DOCUMENT RESUME

ED 420 573 SO 028 708

AUTHOR	Lavy, Victor
TITLE	The Effect of Class Heterogeneity on Scholastic Achievement in Israel.
INSTITUTION	Jerusalem Center for Public Affairs, Inc. (Israel).; Foundations of the Milken Families, Jerusalem (Israel).
PUB DATE	1997-06-00
NOTE	84p.
AVAILABLE FROM	Milken Institute for the Study of Educational Systems, Beit Milken, 13 Tel Hai Street, Jerusalem 92107 Israel; telephone: 02-5619281.
PUB TYPE	Reports - Research (143)
LANGUAGE	English, Hebrew
EDRS PRICE	MF01/PC04 Plus Postage.
DESCRIPTORS	*Academic Achievement; Comparative Education; Cross Cultural Training; Cultural Interrelationships; Elementary Education; Foreign Countries; *Heterogeneous Grouping; Homogeneous Grouping; Multicultural Education; School Desegregation; Social Behavior; *Social Class; *Social Integration; Socioeconomic Status
IDENTIFIERS	*Israel

ABSTRACT

This study examines the relationship between class heterogeneity and scholastic performance in Israel's primary schools. The effect of school integration on academic achievement is explained via two main effects: (1) the "peer" effect, namely externalities that are induced by the composition of the teaching and learning environments; and (2) the efficiency effect, which reflects the reduced ability of the teacher to teach and the pupil to learn in a heterogeneous environment. This study focuses on the efficiency effect and estimates an upper bound to the "peer" effect. The raw, strong negative correlation between achievement and class heterogeneity is clearly an artifact of the association between more heterogeneous classes and lower socioeconomic status among pupils. A liberalization in the education system allowing parents more freedom in choosing an elementary school for their child is seen to be positive in terms of schooling efficiency and the implied gains in scholastic achievement. This liberalization effort is weighed against the setback in the efforts to achieve social integration through the education system and the implied, but unmeasured, benefits. (EH)

מכון מילקן לחקר מערכות חינוך
ע מ ו ת ה

The Effect of Class Heterogeneity on Scholastic Achievement in Israel

by

Prof. Victor Lavy

Jerusalem, June 1997

Prof. Daniel J. Elazar
Academic Director

Zvi R. Marom
Director General

A joint project of the Jerusalem Center for Public Affairs and the Milken Family Foundation

פרוייקט משותף של המרכז הירושלמי לענייני ציבור ומדינה וקרן משפחת מילקן-לשעבר אייסס

בית מילקן, רח׳ תל-חי 13, ירושלים 92107 ☎ טל. 02-5619281 פקס. 02-5619112
E-Mail: elazar@vms.huji.ac.il :דואר אלקטרוני

Milken Institute for the Study of Educational Systems
(formerly ISES)

The Milken Institute for the Study of Educational Systems is a joint venture of the Jerusalem Center for Public Affairs and the Milken Family Foundation. The establishment of this new, separately incorporated institute was made possible by a generous grant from the Milken Family Foundation. The aim of the institute is to conduct applied research on educational problems from a systematic perspective in order to deal with the educational needs of the 21st century.

Address: Beit Milken
13 Tel Hai Street
Jerusalem 92107 Israel

Tele: 02-5619281
Fax: 02-5619112
Email: elazar@vms.huji.ac.il

Abstract

In this study we examine the relationship between class-heterogeneity and scholastic performance in Israel's primary schools. The effect of school integration on academic achievement is explained via two main effects. The first is the 'peer' effect, namely externalities that are induced by the composition of teaching and learning environment. The second is the efficiency effect which reflects the reduced ability of the teacher to teach and the pupil to learn in a heterogeneous environment. While most previous studies in Israel concentrated on measuring the externality effect, our study focuses on the efficiency effect and estimate an upper bound to the 'peer effect'. We present a variety of OLS and IV estimates of the effect of class-heterogeneity on the achievement of Israeli elementary school children. The raw strong negative correlation between achievement and class-heterogeneity is clearly an artifact of the association between more heterogeneous classes and lower SES among pupils. The effect of class heterogeneity is also very sensitive to the control for school size and educational resources. Even after controlling for all these important variables the estimated effect of class-heterogeneity is significantly biased downward since the extent of school integration is a policy variable that is not independent of unmeasured determinants of cognitive achievements. IV estimates using the lagged values of class-heterogeneity and class SES consistently show a negative association between class-heterogeneity and student achievement. The class-heterogeneity effect dominates the estimated upper bound of the 'peer' effect, suggesting that school integration has a net negative effect on scholastic achievement. These effects are apparent for both the math and reading scores of 3rd, 4th and 5th graders. The effects we found are very different from those reported in earlier studies of school integration in Israel. The size of the effect reflect significant impact relative to the math and reading comprehension score distributions. Several Israeli local authorities recently began implementing a new policy that will allow parents more freedom in choosing an elementary school for their child instead of the current strictly enforced school zone rules. Our results suggest that the increased school homogeneity that will necessarily follow from this liberalization in the education system will have a positive side to it in terms of schooling efficiency and the implied gains in scholastic achievement. This gain could lead to impressive improvements in Israeli test scores but this must be weighed against the setback in the efforts to achieve social integration through the education system and the implied, but unmeasured, benefits.

Thanks go to Joshua Angrist and Ruth Klinov for helpful discussions along the way and to participants of the applied workshop on the department of economics at the Hebrew University for useful comments. Nathalie Aflalo, Dan Navot and Dan Ben-Chour provided expert research assistance. The author bears sole responsibility for the content of this paper.

1. Introduction

In this study we examine the relationship between class-heterogeneity and scholastic performance in Israel's primary schools. Ever since the Coleman Report (1966) concluded that the most significant determinant of a student's scholastic achievement, apart from his ability, was the ability of his classmates, school heterogeneity and integration have been major issues for parents, school authorities, and social scientists. In the case of Israel, itself a melting pot of immigrants from many different countries, the integration of schools was perceived by its proponents, as well as the public at large, to be not only a priority for improving scholastic achievement in schools, but also as a way to enhance social and cultural integration among the diverse ethnic groups in the new Jewish state. The implementation of integration policy in Israel since the early 1960's has reflected two basic values of Israeli society: equal opportunity on the one hand and the assimilation of Jews from the diaspora on the other. The integration policy had to deal with scholastic and social gaps which were defined principally, but not exclusively, along ethnic lines between immigrants from Asia-Africa and immigrants from Europe-America (Ben-Ari, Sharan and Amir, 1978; Klein and Eshel, 1985).

The main effort to enhance the integration of students in Israeli schools by administrative means was carried out jointly with the 1968 reform of the school structure and the creation of the junior high school (grades 7-9). The zoning of these new schools attempted to integrate students from diverse socioeconomic backgrounds in each junior high school. Early studies attempted to evaluate the effect of this reform by examining the effect of the integration program on educational outcomes in Israel. Over the last two decades the debate over the effectiveness of Israeli schools has focused on the educational consequences of the integration policy.

Though there has never been a centrally administered national program to enhance integration in primary schools, local authorities in many municipalities took the initiative, especially during the 1980's, to change the zoning of schools so as to create a less homogenous student populations. However, in some Israeli cities, parents have recently been allowed more freedom in choosing a primary school for their children, a move that will reduce the heterogeneity of school populations achieved during the previous decade. This study aims at providing updated empirical evidence that can be used to assess the likely impact of a more homogenous learning environment on scholastic achievement in elementary schools in Israel. While previous studies focused on the effect of externalities related to the 'peer effect' of integration policy, this study will concentrate on the effect of an integration policy on the efficiency of instruction and learning. The results should be useful to both policy makers and parents in examining the pros and cons of parental free choice versus the strict assignment of children to a specific primary school.

The accumulated research on this topic has yet to indicate clear conclusions or policy recommendations. Dar and Resh (1988) provide a summary of 12 studies on the effect of integration policy in Israel. The results are not conclusive and there exist conflicting conclusions as to the effectiveness of integration policy. The study of the effect of class-heterogeneity on student achievement in Israel has concentrated on the effect of ethnic integration in the school. These studies reveal a positive effect of integration on the scholastic achievement of the weaker group. The effect, though very weak, seems more consistent than the one detected in American studies of school desegregation. Despite the vastness of the literature, American research on this subject has not yet produced unequivocal findings as to the effect of racial composition on achievement, though academic improvement among desegregated black students is generally noted.[1]

The effect of ethnic composition on scholastic achievement in Israel is clearest in studies conducted in a quasi-experimental framework, that is, studies in which a comparison was made between a heterogenous treatment group and a homogenous control group, even when the allocation of students to research groups was not random (Arzi and Amir, 1977; Klein and Eshel, 1980; Katz and Ben-Yochanan, 1988). However, the integration effect was significant only in combination with additional educational interventions, such as activity-oriented instruction (Arzi and Amir, 1977), or instruction in small groups with additional hours of instruction and special attention to weak students (Katz and Ben-Yochanan, 1988).

Most of the other studies dealt with "natural" integration and concentrated on measures of the composition of students as proxies for the level of integration in a class. The most commonly used measure was ethnic composition. Multiple regression revealed that in no study did ethnic composition add more than 4 percent to the explanation of achievement variance beyond the 20-30 percent usually explained by background variables. In most cases ethnic composition contributed even less, and in some cases its contribution approached zero.

The compositional effect in studies of "natural" integration appears strongest when integration was conceptualized as the social composition of the school or class, including both

[1] For example, according to Good and Brophy (1987) the desegregation effort has not improved test scores for blacks or whites. In addition, Farely and Allen (1987) have noted that efforts to integrate schools in large urban areas have not worked. Slavin (1986) showed that high achievers do better for a while if grouped in homogenous classes. However,the removal of these high achievers from other classes hurts the achievement of the remaining students while the initial benefit to the high achievers also disappears over time. Other studies include Coleman et al. (1966), Spady (1973), Gerald and Miller (1975), Bradely and Bradely (1977), Alexander and Cook (1982), Link and Mulligan (1991).

ethnic and socioeconomic dimensions (Smilanski and Shfatia, 1977). While these studies controlled for student background variables - father's ethnic origin and education and family size - their models also included aggregates of the same variables as attributes of the social composition of the school or class. Studies that included only ethnic compositional variables as a measure of integration (by including socioeconomic compositional measures in the regression but without conceptualizing them as components of the integration measure) reported a very small effect of ethnic composition on achievements (Minkevich, Davis, and Bashi, 1980).

Three features are common to all the non-experimental empirical studies on school integration in Israel. The first is the treatment of the various measures of integration as predetermined variables that could be used as legitimate explanatory variables in any schooling outcome regression. As noted above, the extent of integration in junior high schools in Israel is directly manipulated by policy makers, and therefore it might not be independent of unobserved (to the researcher) effects on schooling outcomes. In many primary schools the extent of integration might also be endogenous, determined by decisions of school authorities who take into account factors (observed and unobserved to the researcher) that are closely related to schooling outcomes, such as the past achievements of students or other measures of school quality (for example, the ability of the school principal).

The second common thread linking previous studies on integration is the lack of any control of educational and economic resources in schools. This omission could be important given that school resources in Israel are largely a function of the socioeconomic composition of the students and in middle schools also an explicit function of the variance of the student population. Schools with a high proportion of students from low socioeconomic status (SES) and schools with higher SES variability, are supposed to receive more resources, mainly in the form of teachers and hours of instruction.[2] Therefore, any change in the heterogeneity of students in a school, which may or may not lead to a change in the average composition of the student body, will also lead to a change in resources allocated to the school. Omitting school resources from the school outcome equation can lead to a bias in the estimated effect of student heterogeneity. However, simply controlling for school resources will not solve the problem since the resources are themselves an endogenous input. As will be shown later, the estimated effect of class-heterogeneity is sensitive to the omission of school resources from the schooling outcome equation and to whether the endogeneity of school resources is accounted for in empirical estimation. The last common feature is that all concentrated on the 'peer effect' and neglected the efficiency effect.

The objective of this study is to identify the efficiency effect of an integration policy by estimating the causal effect of class-heterogeneity on the scholastic achievement of students in primary schools in Israel. We use a unique data set that includes all the Jewish public primary schools and provides data on test scores in math and reading for 3rd, 4th and 5th graders from

---

[2] Since 1972, special help granted to deprived Jewish children has been institutionalized by the Ministry of Education. A special administration was set up to allocate extra resources and distribute these according to two indices: an index of educational deprivation (an SES index) and an index of "Polarization" (the within school standard deviation of the SES index), i.e., the extent of heterogeneity in the academic ability of pupils within each school. The most important support takes the form of extra weekly teaching hours; another support enables a school to maintain smaller classes. The relationship between allocated resources in our data (weekly hours of instruction and class size) and their determinants (the SES index, its variability, and school size) yields the expected results. First, low SES and more heterogenous primary schools get more teaching hours per student and have smaller classes. It seems therefore that primary schools are also compensated, like middle schools in Israel, for higher level of heterogeneity. Second, school size is a significant determinant of school resources: larger schools receive fewer teaching hours per students and have larger classes.

a 1991-92 national testing program. The data also includes a measure of class-heterogeneity, the standard deviation of a socioeconomic (SES) index calculated for each school, based on micro student-level data. This data was collected periodically, allowing the computation of the SES index and its variance in 1984 and 1990. We use the lagged 1984 values of the SES index and its standard deviation as instruments to construct instrumental variables for the regressions of test scores on class-heterogeneity and school characteristics.

As noted above, resources are allocated to public schools by a social planner from the Ministry of Education. The endogeneity of school inputs is therefore related to the allocation of public programs, in particular integration policy. The estimation of the effect of public programs or inputs in cases where program allocation or inputs are endogenously determined has been discussed in several recent studies (Rosenzweig and Wolpin, 1986) and several econometric techniques have been suggested to deal with the problem. In this study, we take advantage of the fact that in our sample, the distribution of resources across schools is based on an explicit formula. However, even though the allocation mechanism is known exactly, it explains only 50 percent of the variation in school resources, suggesting that the inter-school variation in actual school inputs is not wholly explained by the observable allocation criterion. Since the unmeasured part of the determinants of school resources may not be orthogonal to unmeasured determinants of the outcome in the school production function, the endogeneity and identification problems are not completely resolved by knowing the 'formal' allocation mechanisms (which are based solely on observables). To rid the analysis of potential endogeneity bias we use as an instrument a variable derived from a unique administrative rule that imposes an upper bound to class size in primary schools. It therefore has an independent effect on school resources, such as class size and weekly hours of instruction per student, and has no independent effect on the outcome variable.[3]

A key school characteristic to control for is school size. It is correlated with the variance of school-composition and heterogeneity and it has an independent effect on scholastic achievement. Thus, omitting school size as a control in an achievement regression can lead, as it does in our case, to a spurious relationship between class-heterogeneity and achievement.

The paper is organized as follows: section 2 reviews the details of the integration program in Israel and empirical evidence from the many studies that have examined its effects on schooling outcomes. Section 3 describes the statistical model and the framework for inference and section 4 discusses the identification strategy of the endogenous variables. Section 5 presents the data and section 6 reports the estimation results. Interpretation and implications of the evidence and their summary are presented in sections 7 and 8.

The findings suggest that an increase in class-heterogeneity induces a significant and substantial decrease in math and reading achievements for 3rd, 4th, and 5th graders. This effect is smaller in schools with low average SES, but its sign reverses only in the very low end of the range of the SES index observed in the data. Treating heterogeneity as a predetermined regressor leads to very different results: the effect of heterogeneity is negative in most cases but in none of the cases is it significantly different from zero. The omission of school resources, especially weekly hours of instruction, from the cognitive achievement equation causes a drop (approximately 25 percent) in the effect of heterogeneity on test scores.

[3] This instrument was first developed and used in Angrist and Lavy (1996) in their study of the effect of class size on cognitive achievement in Israel.

The results are robust to various changes in specification and choice of sample. For example, the use of micro student data (3rd graders) yields the same results as the data on class means. Limiting the sample to the largest six municipalities in Israel does not influence the results, suggesting the absence of regional variation in the causal relationship between class-heterogeneity and student achievements. Since class-heterogeneity in the large municipalities was more likely to be affected by school authorities than in the rest of the sample, this result suggests that the effect of class-heterogeneity in a manipulated policy environment is not different from the effect of "natural heterogeneity" observed in our sample.

In addition to being of general interest as a contribution to the Israeli and international class-heterogeneity debate, these findings have immediate relevance for policy making in view of the recent trend in Israel toward increased parental choice of primary schools in certain large municipalities, most notably Tel-Aviv, which has re-ignited the integration debate. Since many other municipalities are currently contemplating the abolishment of strict elementary school zoning, which will most likely lead to more homogenous schools. The supporters of integration are launching a campaign to halt this trend.

2. The Integration Policy in Israel: Facts and Evidence

Beginning in the early sixties and continuing into the seventies and eighties, a policy of integration was implemented in the education system in Israel. By integration we refer to the policy of placing students of Oriental ethnic origin, characterized by lower achievement oriented environments and lower socioeconomic status, together with students of Western ethnic background who are characterized by higher achievement-oriented environments and higher socioeconomic status. The integration policy in Israel was viewed not merely as a means of increasing the formal equality of education, but also as a stratagem for achieving a higher level of national integration. It was the most important device for realizing the ultimate goal of "intermingling of the exiles." The pursuit of these dual objectives required a large investment of pedagogical resources. Whether directed to promote disadvantaged students or to cope with heterogenous classrooms, this input may be an important factor in augmenting the effects of educational integration.

The extent of integration at the junior high school level was never documented and made public, even though the integration policy was an openly stated objective at the junior high school level. Similarly, the level of integration in primary schools, which is the focus of this study, was also never documented and made public. However, in 1984 and 1990, the Ministry of Education updated the SES index it had been using since the mid 1970's to allocate resources to schools and published, for the first time, the variance of the SES index (within school variation). The standard deviation of the updated SES index was termed the 'polarization index' and was viewed as a good measure of the extent of social and economic integration at the school level.[4]

We have used the polarization index to examine the distribution of class-heterogeneity and its change over the 1984-90 period in Israeli primary schools. Again we would like to re-

[4]Klinov (1993) was the first to use the polarization index to examine the resource allocation process of the Ministry of Education. The data obtained from the Ministry of Education by the author included only state non-religious schools, a constraint which was relaxed in this study.

emphasize that this analysis is limited to the state Jewish schools. The sample of more than 1100 schools constitutes almost all of the student population of state Jewish schools.

Despite the short time interval involved (1984 to 1990) important differences in the dynamics of the integration process are revealed. For the country as whole the extent of integration declined over the period under study. As noted in Klinov (1993), the reduction in average heterogeneity is a result of the decline in the size of the extremely low SES groups (particularly those with over four siblings and/or fathers having fewer than ten years of schooling). But over and above the changing demographic structure, the decline in polarization could also arise from a trend toward less social integration. This trend could be the outcome of efforts by the public to countervail the government's pressure toward integration (through geographic mobility and the creation of independent and private schools) or of less effort on the part of some local school authorities to integrate schools. Some light can be shed on this issue by examining the integration dynamics in large municipalities (for example, Tel-Aviv[5], which liberalized the zoning of primary schools in recent years) versus the rest of the sample. Very little difference was found among the rates of change of polarization in various regions of the country, except within Tel-Aviv. The level of integration in northern Tel-Aviv declined at a rate which was 75 percent higher than in southern Tel-Aviv. Could this be a result of 'fatigue' in the integration effort in Tel-Aviv or is it the result of the trend toward liberalization in primary school zoning? Our data cannot be used to address this question.

---

[5]We have chosen Tel-Aviv both because it is relatively easy to divide the sample of schools into high and low SES schools using geographical boundaries (not feasible in Jerusalem or Haifa) and because a conscious effort was made in Tel-Aviv to integrate students from the southern and northern neighborhoods.

Before turning to a review of the literature on the effect of class-heterogeneity on student achievement we should also note that in Tel-Aviv the percentage difference in heterogeneity between students in southern and northern schools (about 20 percent) is much lower than the difference in the SES index (more than 500 percent). The relative magnitudes of the gaps in 'quality' of student and class-heterogeneity are central to the issue of school integration. Regardless of the direction of the effect of integration on cognitive achievements, i.e., whether it improves or hampers the performance of students, it is obvious that an important gap in terms of magnitude is 'socioeconomic'. We will return to this issue when discussing the policy implications of our findings.

Evidence on the Effect of School Integration in Israel

During the seventies a large amount of research attempted to assess the effect of "natural" and intentional integration on academic outcomes in Israel.[6] Lewy and Chen (1974) studied the effect of "natural" integration upon the scholastic progress of students from grade four to grade six in a national sample of 100 classes from 69 schools. Integration was measured by the percentage of Oriental students in the class. On the whole, ethnic composition made a small, though not insignificant contribution to the explanation of achievement. The same data were further analyzed by Lewy (1977) who concentrated more specifically on compositional effects. Based on his finding that the compositional variables added between one and four percent to the

[6]We draw information from the survey in Dar and Resh (1988) regarding studies published up to the mid 1980's.

explained variance, Lewy concluded that classroom composition was not a significant factor in explaining student progress.

In a survey of Israeli elementary school education in 1972/3, Minkevich, Davis, and Bashi (1980) measured the effects of "natural" ethnic integration on achievements in reading, mathematics, English, geography, and science which were assessed cross-sectionally in grades one, two, four and six in a comprehensive national sample of 17,700 students in 614 classes. Ethnic composition added virtually nothing to the explanation of variance in achievement. Hence the researchers concluded that "data on ethnic composition do not improve the prediction of mean class achievement." They further concluded that "integration as carried out in the school does not generally affect students' acumen." No attempt was made in this study or in the two earlier ones, to assess interaction between the student's socio-economic and cultural background and the class's socioeconomic and ethnic composition. Smilanski and Shfatia (1977) confined their analysis to the classroom level and distinguished clearly between personal and classroom characteristics. Their study was designed to assess the effect of "natural" integration on achievement. Integration was measured by classroom composition (percentage of Westerners and percentage of fathers with post-elementary education). The contribution of classroom composition to the explained variance of achievement was small (2.5 percent out of a total 21.6 percent) but significant. Arzi and Zamir (1977) compared the cognitive achievements of a sample of integrated schools to a comparable sample of homogenous schools. Their results indicated a modest advantage for disadvantaged students in an heterogeneous educational environment.

Resh and Dar (1990) is a more recent study that examined the effect of integration in a national sample of junior high schools. The effect of school heterogeneity was found to be very weak and insignificant. The extent of segregation had a negative (but not significantly different from zero) effect on reading and science achievement. As in all the previous studies, the extent of heterogeneity was treated as exogenous and did not control for school and educational resources.

Few studies have specifically investigated composition effects while distinguishing between high and low ability. Dar (1980) compared student performance in a sample of about 700 tenth, eleventh and twelfth graders who attended six kibbutz high schools during 1971-72. In this study a quasi-experimental comparison was made between two categories: heterogenous classes versus homogenous classes of low and high ability levels. This study revealed a small advantage for heterogeneity. However, the advantage for heterogeneity was concentrated in the lower half of the ability distribution. Low ability students in heterogenous classes have a significant scholastic advantage over similar students in low ability homogenous classes. On the other hand, there is real disadvantage to heterogeneity for the high ability students. Dar and Resh (1981) examined a nation-wide sample which was both ethnically and socio-economically heterogenous, and performed a re-analysis of data on junior highs collected in 1972-74 and used in Chen, Lewy and Adler (1978). Their study also indicated that the less endowed students were found to be more sensitive to the level of heterogeneity in their class.

We turn next to a review of those studies which examined intentional integration. The first such study, the "Nachlaot project" in Jerusalem, provided an opportunity to assess planned integration by variation in method of instruction in a quasi-experimental design (Klein and Eshel, 1980). Students in integrated and non-integrated classrooms were observed over a five year period, from first to sixth grade. Comparisons of achievement in math and reading comprehension revealed no difference between low SES students in integrated classes and those in non-integrated classes. The level of achievement of high SES students was slightly lower in integrated classes. However, students in integrated classes in which non-conventional teaching methods were used had significantly higher test scores than students taught using conventional methods, whether they were in an integrated or non-integrated class. The authors concluded that integration was effective only when accompanied by appropriate pedagogical input.

The integration policy in Israel was accompanied by a reform that led to the creation of the junior high school as a separate educational and organizational unit for grades seven to nine. The new junior high schools increased ethnic heterogeneity relative to the former elementary schools. The junior high school study by Chen, Lewy and Adler (1978) is the most comprehensive of the Israeli studies which examined the outcomes of integration and is the only one conducted on a large-scale at the post-elementary level. The scholastic progress of students in junior high schools was measured over the period 1971-73 and compared to control samples from non-reformed schools and was composed of seventh and eighth graders from primary schools and ninth graders from high schools. The comparison of achievements between junior high students and eighth graders in elementary schools and ninth graders in high school, suggested only a slight and insignificant disadvantage for the junior high's eighth grade. The authors concluded that the new and more integrated junior high schools had not lead to improved scholastic achievement.

Katz and Ben-Yochanan (1988) is an evaluation of an integration project that started in 1984/85. First grade students of Oriental ethnic origin with lower SES levels were placed together with students of Western background with higher SES levels. The control group included schools in the area from which the Oriental students were sampled. The experiment also included educational and integrational interventions in the experimental group, which were considered to be conducive to improved academic achievement. The study concluded that the lower level students that were placed in the integrated school achieved significantly higher scores on a reading comprehension test than their control group peers. At the same time, the higher level students maintained a superior level of achievement according to this measure despite the efforts made to close the gap between higher and lower level students participating in the integration project. This evidence, however, cannot be used to analyze the effect of class heterogeneity on achievement due to the additional educational interventions included in the experiment. The achievement increment may be attributed to the resources utilized for the benefit of the lower level students participating in the integration project. The teaching methodology employed in the integration project (small learning groups and cooperative learning, vocabulary and auditory enrichment, as well as 'big brother' tutors who provided the lower status students with extra help in preparing their homework and preparing for tests) was thought to have contributed to the academic improvement of the experimental group.

To summarize, the studies of ethnic integration in Israel have shown a positive, if weak, correlation between the quality of the learning environment (as measured by the percentage of Western origin students in the classroom) and achievement (principally in math) primarily in the elementary school. The effect is more marked in studies conducted with a quasi-experimental design which compared integrated and segregated groups. Most of the studies, however, dealt with "natural" integration and applied one of two methods in assessing its effects: One approach attempts to estimate the net addition to the explained variance contributed by ethnic composition, having taken into account personal scholastic and socio-economic data and even socio-economic composition. Dar and Resh (1988) claim in their review that due to high collinearity between SES and ethnic composition, this approach leaves virtually nothing to be explained by the latter and researchers erroneously concluded that ethnic integration has no effect on scholastic achievement. In fact, all that can be said is that the net ethnic component has no effect beyond that of the socio-economic component. The second approach, which distinguishes between personal and compositional variables, attempts to assess the effect of both inter-correlated compositional components while controlling for personal variables. This approach reveals the positive effect of integration more clearly, with the ethnic and/or socioeconomic composition of the classroom appearing as a significant predictor of scholastic achievement, albeit a much weaker one than personal background variables.

From our point of view, all of the studies surveyed above have failed to account for the endogeneity of school-heterogeneity (except for the experimental studies), did not control for school educational resources, and did not control for school size which is highly correlated both with school-heterogeneity and school resources and has an independent effect on schooling outcome. All of these factors could cause a bias in the estimated effect of class-heterogeneity on academic achievement.

Finally, all of the non-experimental studies concentrated in measuring the effect of externalities associated with school integration but did not pay attention to its potential effect on teaching and learning efficiency.

3. Estimation Framework

The education literature provides several explanations and mechanisms through which school integration can affect the learning outcome of an individual. In terms of the economic terminology these can be summarized and grouped into two main categories: externalities and efficiency of instruction and learning. The main justification given for school integration is in terms of externalities that arise through the interaction of high and low achievers. Three kinds of externalities can be identified: First, there are externalities that are induced by the quality of learning interaction, which is determined by the level, speed of progress and content of the studies, accessibility of role models among the students and the learning demands defined by the teachers (Oakes, 1985, Lee and Bryk, 1988). Second, there are externalities related to motivational factors which are derived from social-psychological theories and from a perception of the class as the reference group. With regard to learning behavior, the perceptions relate to norms about the value of schooling and derived expectations about learning effort. The composition of students determines the learning norms and the learning climate in class, and generate group pressure on individuals to behave according to these norms (Oakes; 1985). Finally, there are externalities that are related to the development of learning and social self-image, including perceptions of the future value of learning and investment in quality and quantity of schooling. These externalities could lead to a positive effect of school integration on the learning outcome of low achievers and generate gains that can more than offset the losses of high achievers.

Opponents of integration counter the above with an argument related to the deductive match of the student to the learning material, its complexity and the pace it is presented. The ability of the teacher to assure such a deductive match is enhanced if the learning group is more homogenous in its learning skills and level of learning motivation (Bar and Dreeben; 1983). Therefore an integration policy that creates a more heterogenous learning environment reduces the efficiency of teaching and learning by limiting the ability to achieve the deductive match to most of the students.

To best measure the effect of externalities requires information about student ability and that of his peers. To measure the efficiency effect requires a measure of class-heterogeneity. In this study we lack micro information on student ability, but we have a good measure of school heterogeneity. We therefore focus the empirical work on measuring the efficiency loss involved with more heterogeneous learning environment. However, we also estimate an upper bound for the net 'peer' effect which allows us to determine the lower bound of the net effect of school integration.

We use a statistical model for individual pupil's test scores in order to capture any causal relationships of interest and develop a framework for inference. For the *i*th student in class *c* and school *s*, let

$$y_{isc} = X_s'\beta + a_{isc}\mu + p_{sc}\alpha + h_{sc}\delta + \epsilon_{isc} \qquad (1)$$

where y_{isc} is pupil i's score, X_s is a vector of school characteristics (for example school size), and school inputs. We consider two main inputs, total weekly hours of instruction per student and class size. a_{isc} is pupil i's ability (his SES index), p_{sc} is the class average for ability (class average SES index) and h_{sc} is the level of heterogeneity of students in class c in school s. ϵ_{isc} is an i.i.d. random error term specific to pupils.[7] The coefficients α and δ are the parameters of primary interest. The first is a measure of the effect of externalities (peer effect), while the second is a measure of the efficiency effect of class-heterogeneity.

Equation (1) is intended to describe the potential outcomes of students under alternative values of p_{sc} and h_{sc}, when X_s and a_{isc} are held fixed.[8] Note, however, that since p_{sc} and h_{sc} are not randomly assigned, they are likely to be correlated with potential outcomes (i.e. the error term in [1]). On the other hand, since (1) is defined as a causal response function with X_s and a_{isc} held fixed, the vector of exogenous regressors is uncorrelated with the error components by construction.

Equation (1) is cast at the individual level though most of the literature on class heterogeneity treats the class rather than the student as the unit of analysis. Since class-heterogeneity is naturally fixed within classes and student test scores are correlated within

---

[7]Angrist and Lavy, 1996, used a model that allows for the fact that students' test scores are likely to be correlated within classes, and, to a lesser extent, within schools. They decompose the error term to a random class component, to a random school component and to a term specific to each pupil. The three error components, assumed to be mutually orthogonal, were introduced to parameterize possible within-school or within-class correlation. However, the correlation of class averages within schools did not have a noticeable impact on the standard errors of coefficient estimates, leading to standard errors 5-10 percent larger than conventional formulas. We therefore opted in this study to use the conventional standard errors.

[8]Although equation (1) is linear, linearity of the true causal response function is not essential for estimates of α and δ to have a valid causal interpretation (see Angrist and Imbens, 1995).

classes, little is likely to be lost in statistical precision from this aggregation. Moreover, in the case of the test results for 4th and 5th graders, we have no option other than a class-level analysis since the micro-level data has been lost or destroyed. To make the analyses comparable, we also aggregated the 1992 data on 3rd graders to the class level.

The class-level estimating equations have the form:

$$\bar{y}_{sc} = X_s'\beta + p_{sc}(\mu+\alpha) + h_{sc}\delta + \bar{\epsilon}_{sc} \tag{2}$$

where overbars denote averages. Estimation of equation (2) yields an estimate of the efficiency effect of class-heterogeneity, δ, but not of the externalities effect of school integration, α. The coefficient of p_{sc} in equation (2) is the sum of the effect of the pupil's ability and that of his peers. Therefore this sum can serve as an upper bound estimate of the peer effect which can be compared to the efficiency effect in order to assess the net effect of school integration on scholastic achievement.

The empirical results consist of ordinary least squares (OLS) and instrumental variable (IV) estimates of (2). The instruments are generated from the lagged values of the SES and class heterogeneity variables and from a special administrative rule that imposes a maximum of 40 students per class in primary schools in Israel.

4. Identification of the endogenous variables

School heterogeneity

The level of heterogeneity within a class is related to the heterogeneity of the population in the zoning area of the school. The zoning area of the school is determined by the local government's educational authority. The size of the zone, the opening of new schools and the closure of existing schools as well as the specific size of each school, are all functions of demographic changes in surrounding neighborhoods. Large schools will naturally draw students from a large zone. The population of a randomly drawn school zone will naturally be less homogenous the larger the zone. Therefore we should expect to have a positive correlation between the heterogeneity of the student population and the size of the school. However, the boundaries of a school zone are not drawn randomly. In fact, they are often manipulated to increase or decrease the integration level at the school. Therefore the nature of the correlation between school size and school heterogeneity cannot be predicted a-priori in an environment manipulated by policy decisions. Indeed, our data shows a negative correlation (-0.15 with standard error of 0.0001) between school size and school heterogeneity. However, the size of the school is not only correlated with school heterogeneity, but may also have a direct effect on student performance, due, for example, to economies of scale in the use of educational and economic resources. This means that school size belongs in an equation that examines the effect of school heterogeneity on student scholastic achievement.

The heterogeneity of schools may be affected by the policies of the local school authority which are intended to integrate students from different socio-economic backgrounds. As noted in section 2, effort and resources have been directed primarily to the integration of junior high schools, so that the degree of homogeneity in primary schools is mainly a result of 'natural' integration in the school zone. However, there were many instances in which the zone of a primary school was defined with the intent of achieving a higher level of integration or to strengthen a weak school. In both cases the level of heterogeneity may be endogenously determined since it is being correlated with the error term of the cognitive achievement equation.

In principle, heterogeneity can also be affected by parents' selection of schools. If parents take into account the level and heterogeneity of students in the choice of school for their child, then the endogeneity of heterogeneity again becomes an issue. In Israel parents do not have much discretion in choosing schools at the primary and middle school level. Strict zoning rules determine which school a child can attend and the only decision left to the parents is their choice of domicile.[9] Of course one cannot preclude the possibility that choice of schooling is a criteria for choosing where to live. One cannot, therefore, make the assumption that the level of heterogeneity in a school is independent of unobserved factors that affect student performance, which are embedded in the error term of the cognitive achievement equation.

Another important factor affecting school heterogeneity in 1990-91 is the large inflow of immigrants in second half of the 1980's. The location of immigrants to Israel are largely affected by the preferences of government authorities. However, the immigrants do exercise some discretion about where to live in Israel. The government placement policy and the choices made by the immigrants are most likely correlated with characteristics of the localities, in particular the quality of social services, including schooling, and with the heterogeneity of the population.

Some or all of the above explanations for the potential endogeneity of heterogeneity in schools can lead to a bias in the effect of this variable if it is used as a regressor in an OLS regression of equation (2). To break the correlation between this regressor and the error term in equation (2) we will use an instrumental variable procedure. The instrument we propose is

[9]Strictly speaking, each zone has two schools, a religious and a non-religious state school, and parents can choose one of the two. In most cases the religious status of the family determines this choice. Very rarely do families 'cross' this line.

lagged values of the level of heterogeneity in school. Lagged values are appropriate instruments if they are not correlated with the error term, $\bar{\epsilon}_{sc}$, of equation (2). This requirement can be violated in our case if $\bar{\epsilon}_{sc}$ is serially correlated since h_{sc}, as noted above, is affected, through an integration policy or parental choices, by lagged values of $\bar{y}_{sc}$, or with unobserved heterogeneity correlated with lagged values of $\bar{y}_{sc}$. We cannot preclude the possibility that $\bar{\epsilon}_{sc}$ is serially correlated. However, if the lagged values used are far enough in the past, the potential risk for a correlation between the instrument and the contemporaneous $\bar{\epsilon}_{sc}$ is minimized. We therefore use as an instrument the earliest h_{sc} values available. In our data they are the 1984 school level measures of the student body heterogeneity. Since the dependent variable in equation (2) is a June 1991 measure, it is very likely that $cov(\bar{\epsilon}_{sc,1991},\bar{\epsilon}_{sc,1984})$ is very small. If the endogeneity derives from decisions made in recent years by school authorities or parents, as suggested above, then we can net out from the contemporaneous variance of heterogeneity the share arising from these recent policy and parental choice decisions. This can be done by keeping only that share of the variance of the contemporaneous heterogeneity that is correlated with its 1984 value. Again, the assumption is that the lagged value of heterogeneity is not correlated with recent policy decisions or parental choices since it preceded them in time.[10] The correlation between the 1991 and the 1984 value of the measure of school heterogeneity is 0.28.

Since any policy-induced changes in school-heterogeneity might lead to simultaneous changes in the school mean SES index, this latter variable should also be treated as endogenous.

[10]See Chamberlain, 1984; Holtz-Eakin, Newey and Rosen, 1988 for a discussion of the use of lagged values of the endogenous variables as instruments.

We use lagged (1984) SES values as instruments for their contemporaneous values following the same logic that rationalized the use of this same strategy in the case of class-heterogeneity.

A final element in our identification strategy is to include locality (city, town, smaller municipalities) fixed effects in equation (2). Since a school authority policy towards integration may affect all schools in the locality, this effect can be absorbed by a locality dummy variables in equation (2). Such a within locality estimation strategy will also take care of any endogeneity in school heterogeneity due to choices related to the locality characteristics. For example, choices made by parents, including immigrants, about their place of residence, or by the government about placement of immigrants.

School resources

As noted above, class size and hours of instruction per student may be endogenously determined. To solve this problem, we exploit a unique feature of the process by which class size is determined in Israel. In 1969 an agreement between the teachers union and the Ministry of Education determined that class size in Israeli primary schools will not exceed 40. Therefore, this rule determines the division of enrollment cohorts into classes in Israeli public schools today and it generates a source of variation in class size that can be used as an instrument for class size or hours of instruction per student. To see how this variation comes about in practice, note that according to this rule, class size increases one-for-one with enrollment until 40 pupils are enrolled, but when 41 students are enrolled, there will be a sharp drop in class size, to an average of 20.5 pupils. Similarly, when 80 pupils are enrolled, the average class size will again be 40, but when 81 pupils are enrolled the average class size drops to 27. Of course, this rule is not the only source of variation in Israeli class sizes, and average class size is generally smaller than what would be predicted by a strict application of this rule. But Israeli classes are large and the ceiling of 40 students per class is a real constraint faced by many school principals. Both the median and the median Israeli class sizes for Jewish students in our data are around 31 pupils, with 25% of classes having more than 35 pupils. A regression of actual class size at mid-year on predicted class-size using beginning-of-the-year enrollment data and the maximum 40 students in a class rule explains about half the variation in class size in each grade (in a population of about 2,000 classes per grade).

The class size function derived from this rule can be stated formally as follows. Let e_s denote beginning-of-the-year enrollment in school s in a given grade and let f_{sc} denote the class size assigned to class c in school s for that grade. Assuming cohorts are divided into classes of equal size, we have

$$f_{sc} = e_s/(\text{int}((e_s-1)/40)+1), \qquad (3)$$

where, for any positive number, a, int(a) is the largest integer less than or equal to a. Equation (3) captures the fact that the rule allows enrollment cohorts of 1-40 to be grouped in a single class, but enrollment cohorts of 41-80 are split into two classes of average size 20.5-40, enrollment cohorts of 81-120 are split into three classes of average size 27-40, and so on. Of course, any single class necessarily includes an integer number of students, so that an enrollment cohort of 121 would likely be divided into 3 classes of 30 pupils and one class of 31 pupils. In this paper, we use the class-size function induced by the 40 max rule to construct instrumental variables estimates of class-size or hours of instruction per student effects in regressions of test scores on class composition, class heterogeneity and class size or hours of instruction. The class-size function and the instruments derived from it are a function of the size of enrollment cohorts. However, these functions are non-linear and non-monotonic and we can therefore control for a wide range of secular cohort-size effects when using the class-size function as an instrument. It turns out that in practice, there is not always a need to do this since we also control for school size. Angrist and Lavy (1976) show that the class-size function derived from this rule provides an unusually credible source of exogenous variation that can be used to identify the causal effect of class size on student achievement.

5. Data and Descriptive Statistics

The achievement data used in this study came from a short-lived national testing program in Israeli elementary schools. In June of 1991, toward the end of the school year, all 4th and 5th graders were given achievement tests designed to measure mathematics and (Hebrew) reading skills. The tests are described and the results summarized in National Center for Education Feedback (1991). The scores used here consist of a composite constructed from some of the basic and all of the more advanced questions in the test, divided by the number of questions in the composite score, so that the score is scaled from 1-100. This composite is commonly used in Israeli discussions of the test results.[11] As part of the same program, similar tests were given to 3rd graders in June 1992. The June 1992 tests are described in National Center for Education

---

[11]The 4th grade tests included 45 math questions and 57 reading questions. The 5th grade tests included 48 math questions and 60 reading questions. Among these, 15 questions were considered basic and the remainder more advanced.

Feedback (1993).[12] The achievement tests generated considerable public controversy in Israel because of the lower than anticipated scores, especially in 1991, and because of vast regional difference in outcomes. Following the 1992 exams, the national testing program was abandoned.

The analysis begins by linking average math and reading scores for each class with data on school characteristics from other sources. The details of this link are described in the data appendix. Briefly, the linked data sets contain information on the population of schools covered by the Central Bureau of Statistics (1991, 1993) census of schools. These are annual reports on all educational institutions made at the beginning of the school year (in September), based on reports from school authorities to the Israeli Ministry of Education and supplemented by Central Bureau of Statistics data as needed. Our measure of cohort and school size are the beginning-of-the-year enrollment taken from the computerized files underlying these reports and the classes in the schools covered by the reports define our study population. We use the Central Bureau of Statistics files as a reference population because the data from other sources was not collected at exactly the same time and because the reference population tells us how complete the testing program was. In fact, test scores are available for over 96 percent of the classes in the reference population, although due to absences not every pupil in every class was tested.

[12]The 1992 exams included 40 math questions, of which 20 were considered basic. The math composite score included 10 of the basic questions plus 20 of the more advanced questions. The reading exams included 44 questions, of which 20 were considered basic. The reading composite included 10 of the basic reading questions plus all of the more advanced questions.

The unit of observation in the linked data sets and for our statistical analysis is the class.[13] The data sets include information on average test scores in each class, class size, school average of weekly hours of instruction per student, beginning-of-the-year enrollment in the school (all grades and classes), a town identifier and a school-level index of students' socioeconomic status (SES) and an index of its standard deviation.[14] The latter variable is used by the Ministry of Education as a measure (index) of heterogeneity of the student body and as an indicator of the extent of school integration. We therefore use it as the measure of school heterogeneity. Also included are variables identifying the ethnic character of the school (Jewish, Arab, other minority group) and the administrative affiliation of the school (public-secular, public-religious, private-religious). Schools in Israel are segregated along religious lines, and, even within the Jewish public school system, there are separate administrative divisions and separate curricula for secular and religious public schools. Our study is limited to Jewish pupils in the public school system. This group accounts for about 70 percent of the school children in Israel. We exclude all the rest because the measures of SES and heterogeneity were computed only for the public Jewish schools (about 1100 in number). The SES index is a key control

[13]Micro data on students are available for 3rd graders in 1992. For comparability with the 1991 data, we aggregated the 1992 micro data up to the class level.

[14]This is the index discussed by Algrabi (1975) and used by the Ministry of Education to allocate supplementary hours of instruction and other school resources. It is a function of pupils' fathers' education and continent of birth and family size. The index is recorded in our data as the fraction of students in the school who come from what is defined (using index characteristics) as a disadvantaged background.

variable in our analysis because it is correlated with class size and hours of instruction, with school size and with the measure of heterogeneity.

Table 1 provides descriptive statistics for the population of over 2,000 classes (about 56,000 pupils) in each grade. The distribution of test scores, also shown in the table, refers to the distribution of *average* scores in each class.

One important feature of the descriptive statistics is that mean scores are much higher, and the within and overall standard deviations are much lower, for 3rd graders than for the 4th and 5th graders. We believe this is due to a systematic preparation effort for the June 1992 test on the part of teachers and school officials, in light of the political fallout resulting from what were felt to be disappointing test results in June 1991 (see the discussion in Lavy, 1995).

We use class size and hours of instruction per week per pupil in each school as alternative control variables in the cognitive achievement equation. We cannot use both of them jointly as controls since we have only one school level variable to use as an instrument. Hours of instruction per student reflects the ratio of total weekly hours of instruction to average class size. The level of hours of instruction per student is determined for each school by the Ministry of Education as a function of a school-level SES (termed as an index of 'need for improvement'). The SES index is computed as a weighted average of parental educational levels, ethnic background and family size. Hours of instruction per student average 1.46 and ranges from about 1.1 to 2.0, corresponding closely to schools with the minimum and the maximum value of the SES index, respectively. Class size average 30-31 and is very similar for 3rd, 4th and 5th grade. Cohort size average 74-77 for the three grades and school size average 504 students. The mean of the 1990 SES index is .17 (.23 in 1984) and the mean of the 1990 integration index is .23. (.45 in 1984). The large decline in heterogeneity most likely reflects demographic changes as well as reduced government effort to enhance school integration and a relaxation of the strict school zoning rules. Actually the 1990 integration index have zero values for about 200 observations. We will examine the sensitivity of our results to the inclusion of these schools in the analysis.

6. Empirical Results

Treating class-heterogeneity as an exogenous regressor

OLS estimation with no control variables showed a strong negative correlation between class heterogeneity and achievement. However, this association largely disappears, and in all cases it is not significantly different from zero, once the pupils's SES index is added into the equations as a control. Once we add as controls school size and a dummy variable for the religious schools, this association is strengthened somewhat. Adding to the equations town dummies again weakens the association between class heterogeneity and test scores, in some cases it even becomes positive. These findings can be seen in Table 2 which reports coefficients from regressions of the math and reading scores of 4th and 5th graders on the level of class heterogeneity, the SES index, school size, a religious dummy and a set of town fixed effects. For example, in a regression of the average reading scores of 5th graders on class heterogeneity, the class-heterogeneity effect is a precisely estimated -31.205 (column 1 in Table 2); however, when we add the SES index as a control variable, the estimated class-heterogeneity effect declines to -6.711 with a standard error of 1.982. In the regression of math test scores of 5th graders, the coefficient of heterogeneity declines from -28.877 (se=2.322) to -1.780 (se=2.589) when the SES index is added to the equation. The results for 4th graders in columns 9-10 and 13-14 reveal the same results, namely that class heterogeneity has no effect on math and reading test scores once we control for the socio-economic background of the students. When locality dummies are added to the regression, the class-heterogeneity effect in all the 4th grade regressions (reading and math) in Table 2 decreases practically to zero, but in 5th grade it remains negative though marginally significant.

The First Stage Estimates

Two potential endogenous variables need exogenous identifying variation. These are the mean and variance of the school SES index, and, in some specifications, their interaction as well. The identifying instruments used are the lagged values of the SES and its variance and their interaction in the relevant specifications. We estimated two basic models, with and without the interaction of these two variables.

Table 3 presents reduced form estimates for models without the interaction between the SES index and class-heterogeneity. The table includes both the 4th and 5th grade results for reading and math test scores. The heterogeneity and SES index first stage regressions are identical for the reading and math models since the control variables (school size, a religious dummy variable and locality fixed effects) and the instruments (lagged values of SES index and the heterogeneity measure) do not vary by the subject of the test.

The reduced form estimates of the equations for the test scores of 4th and 5th graders show a clear negative association between the 1984 class-heterogeneity and SES index and math and reading test scores. For example, the effect of 1984 heterogeneity on reading test score of

4th graders (column 1 of Table 3) is -3.365 points (s.e.=1.591). The respective coefficient for math is -3.872 (s.e=1.820, column 2 of Table 3). The reduced form estimates for 5th graders show a similar association, in size and precision, between the two instrumental variables and reading and math scores. However, the effect of lagged heterogeneity on the reading scores is lower and very imprecisely estimated.

Another interesting result in Table 3 is the reduced form equations for the class-heterogeneity and SES index. The six-year lagged values of these variables are positively associated with the contemporaneous heterogeneity measure and with relatively low standard errors. For example, the estimated effect of the 1984 heterogeneity value on its 1990 value is .117 (column 3 of Table 3) and it has a very low standard error (.018). Similar positive and precisely estimated association between the SES index and its lagged values is presented in columns 4 and 8 of Table 3. It should also be noted that school-size is highly negatively correlated with heterogeneity and the SES index with t-values of more than 5. The religious schools have, on average, a lower SES index and a higher level of student heterogeneity than the non-religious schools. The coefficients of the religious dummy is precisely estimated in all the regressions presented in Table 3. The instruments explain a large share of the variance of the SES index and a lower share of the heterogeneity measure. The total share of variance explained in these reduced form equations are .19 and .50 for school heterogeneity and the SES index, respectively.

These reduced form estimates can be used to produce simple Wald-type IV estimates of the effect of heterogeneity on test scores. For example, dividing the test score effects in column 1 of Table 3 by the class-heterogeneity effect in column 3 leads to an estimated class-heterogeneity effect on 4th grade reading scores of -3.4/.117 = -28.7, more than six times larger than the OLS estimate of column 4 in Table 3. The Wald-type estimate of the effect of heterogeneity on 4th grade math test score is -3.9/.117 = 33.1, almost 8 times larger than the OLS estimate. Similar results emerge for 5th grade math test scores. Comparing the Wald-type estimators of the SES index to the OLS estimates suggests much smaller gaps between the two. This is an indication that IV estimates of the SES index are not going to be very different from the OLS estimates.

The Second Stage Estimates for 4th and 5th Graders

Tables 4 presents a variety of 2SLS estimates for 4th and 5th graders. For each of the two subjects, reading comprehension and math, we present two regressions. The first is the relevant OLS regression re-estimated with the samples used in Tables 4, the second equation is a 2SLS estimates in which heterogeneity and the SES index are treated as endogenous. All the models include the controls used in the first stage, namely the religious dummy variable, school size and a set of locality fixed effects.

The 2SLS estimates of class-heterogeneity effects using lagged heterogeneity as an instrument in the equations for 4th and 5th graders, reported in Table 4, consistently show a negative effect of class-heterogeneity on achievement for both math and reading scores. The estimate of the effect in a model with heterogeneity instrumented reported in columns 2, 4, 6 and 8 of Tables 4, ranges from -8.661 to -34.971, and it is significantly different from zero in all cases but one. The difference between the 2SLS estimates and the respective OLS is striking, in some cases even opposite in sign. For example, the 2SLS effect of heterogeneity on 4th grade reading test scores is -26.982 with a standard error of 14.366 (column 2 of Table 4) while the OLS coefficient from column (1) is .739 (se=2.029). The effect of heterogeneity on math test scores is larger than on reading and the differences between the two are not trivial, especially in the case of 5th graders.

Two other results should be noted. First, the net effect of the religious dummy is positive and very precisely estimated. Students from religious schools have about 6 points advantage in the reading test and, surprisingly, almost 4 points advantage in math, both in 4th and 5th grades. What seemed initially to be a negative effect of religious schools on test scores is reversed after controlling for the SES level of students and their heterogeneity. Second, what seemed to be an overwhelming positive effect of school size on test scores in the simple OLS regressions totally disappears in the 2SLS reading test score regressions, and halved in size in the math test scores regressions.

Second Stage Estimates For 3rd Graders

In 1992 only 3rd graders were tested in both subjects. As seen in Table 1 the 3rd graders results reveal a significant increase in average test scores and a reduction in variability. This may be the consequence of systematic test preparation efforts on the part of school administrators and officials. Two pieces of evidence support this possibility: First, the public furor raised by the 1991 test results (see e.g., Lavy's [1995] discussion of this point) and second, the official report of the 1992 test results (National Center for Education Feedback, 1993) which noted a number of differences between the two years of the testing program. For example, regular class teachers (as well as the exam proctor employed by an outside contractor) were present when tests were taken in 1992 but not in 1991. A further example was summarized as follows: (page 3, translated from the Hebrew): "During the past year there was an intense and purposeful remedial effort carried out by the Elementary School Division [of the Ministry of Education] in a large number of schools with high failure rates in 1991. Similarly, in light of last year's scores [in 1991], and in anticipation of the forthcoming tests [in 1992], there was an intensive remedial effort on the part of schools, district supervisors, counselors, and others." The evidence presented below for 3rd graders is therefore not perfectly comparable with the 1991 data presented above.

Results for 3rd graders tested in June 1992 are reported in Table 5. Columns 1-3 and 6-8 report OLS estimates of the regression of test scores on the same set of regressors as in Tables 3-4. These results which again show that the OLS coefficient of class heterogeneity is very sensitive to the specification: it changes from being negative and highly significant (columns 1 and 6) to being much smaller and in reading also non-significant. This pattern is identical to the OLS 4th and 5th grade results. The coefficients are -1.884 (se=1.728) for reading in column 3 and -4.999 (se=1.959) for math in column 8.

The IV models presented in columns 4 and 8 are identical to those of the previous table. The results are qualitatively very similar to our previous 1991 results. Most importantly, the IV estimates of the effect of class-heterogeneity on test scores are much larger than the OLS estimates. However, the standard errors, which are larger than those of the 4th and 5th grade, indicate only marginal significance. It is interesting to note that the positive edge of religious schools as indicated by the coefficients of the variable is still kept and significant in the reading but not in the math test scores.

Allowing For The Effect of School Resources

School resources include a wide range of inputs such as hours of instruction, class size, teachers education and experience, materials, computers and more. As discussed above, the allocation of inputs to schools is not independent of measured and not measured school characteristics that affect students test scores. The problem that is raised is how to net out from the actual allocation of inputs that component that is endogenous, namely the part that is correlated with the unobserved determinants of test scores. Two inputs are available in our data, class size and hours of instruction. The first is a class specific variable, the second is a school average measure. Using the class size function derived from the maximum 40 rule we can resolve the identification problem and estimate the effect of class size or of weekly hours of instruction per student. We should note that using hours per student amounts to assuming that the effect of hours and class size on test scores are working only through this ratio and that they do not have any additional effect not captured through this ratio. This is probably a non-realistic assumption but since our primary interest is to estimate the causal effect of heterogeneity and not the effect of the inputs we maintain this assumption.

Tables 6-7 present a variety of OLS and IV regressions for 4th and 5th grades. For the sake of saving space we do not present the first stage regressions of the IV results.[15] All the regressions in these tables include the religious dummy, school size, cohort size and locality fixed effects as controls, although their coefficient estimates are not reported in the tables. Four regressions are reported for each grade and subject. The first two of each set include class size

[15]Angrist and Lavy (1996) presents all the first stage regressions when the class function derived from the max 40 rule is used as an instrument. In this paper the first stage equations also involve the lagged values of the SES and heterogeneity measures.

as a regressor, once as an exogenous variable and the second as an endogenous that needs to be instrumented. In the other two, the class size variable is replaced by weekly hours of instruction per student.

Two main results stand out from Tables 6-7. First, the pattern established from the earlier results, namely that the 2SLS class-heterogeneity effect on test scores is negative, much larger and more precisely estimated than the OLS estimates, is unchanged once we control for class size or hours of instruction per student. Second, the OLS estimates of class size and of hours of instruction highlight the anomaly in the sign of the effect on test scores of these two inputs: class size has a positive effect while hours of instruction has a negative effect.[16] However, when treated as endogenous and instrumented by the class size function these two effects reverse in sign: class size has a negative effect on test scores and hours of instruction a positive effect. In almost all cases these effects are significantly different from zero. For example, the 5th grade class size estimated effect is -.320 (se=.111) in math and -.245 (se=.088) in reading. These class size effects are very similar to those reported in Angrist and Lavy (1997) in their models without town fixed effects. However, the estimates of this paper are larger and more precisely estimated than the Angrist and Lavy (1996) estimates in models

[16]Lavy (1995, 1996) observed that the negative association between hours of instruction and test scores in Israel is largely accounted for by the association between longer weekly hours of instruction per student and lower SES among pupils. A similar point was made earlier in the Coleman Report (1966) regarding the allocation of school resources in general and has been emphasized again recently in the meta-analysis by Hedges, Laine and Greenwald (1994). Angrist and Lavy (1996) also found a positive association between class-size and tests scores in Israel which is partly accounted for by the negative association between class size and the SES index. However, controlling for SES in our data does not completely eliminate the negative association between hours of instruction and reading and math scores.

with town fixed effects. It seems that controlling for class-heterogeneity allow for more precise estimation of the effect of class size. The 5th grade hours of instruction per student effect is 10.512 (se=4.053) in math and 8.026 (se=3.164) in reading. To assess the sensitivity of the measured effect of class heterogeneity on test scores to the control for class size or hours of instruction we summarized the relevant information in Table 8. This summary table assembles from Tables 4-6-7 the IV point estimates of the class heterogeneity variable. When class size is added to the regression, the class-heterogeneity effect is almost unchanged. For example, in the regression of the average reading and math scores of 4th graders, the class-heterogeneity effect on math scores is -31.852 and -30.655 without and with class size included, respectively, with almost identical standard errors.

When hours of instruction per student is included as a control for school resources the effect of class heterogeneity increases significantly: in 4th grade by 19 and 30 percent in math and reading, respectively, in 5th grade by 41 and 120 percent in math and reading, respectively, though the effect on reading has a large standard error. The remarkable sensitivity of the effect of class-heterogeneity on test score to the inclusion/exclusion of the weekly hours of instruction input suggests that these two variables are highly correlated. This correlation is positive, meaning that schools with a more heterogenous student body are endowed with more weekly hours of instruction per student. Therefore the omission of this input from the equation leads to a significant bias in the estimated effect of class-heterogeneity on student achievement.

As noted in the data section the heterogeneity measure for 1990 includes about 200 schools for which this index is zero. This pattern may reflect true values or it may be a result of measurement errors. If indeed the heterogeneity variable is measured with error our instrumenting procedure assures that the estimated effect of heterogeneity is consistent. However, to further prob into the sensitivity of our results to the inclusion of these schools in the sample we re-estimated our models with an alternative specification that included a dummy variable for these schools in all the equations estimated. The results of this specification is practically identical to those reported in Table 8. We also estimated the models with a restricted sample that did not include the 200 observations with the zero value for the school heterogeneity index. Table 9 presents estimates that are comparable, respectively, to those of Table 8, produced with the restricted sample. The evidence in Table 9 is very similar qualitatively to those of Table 8. Quantitatively two differences emerge: first the coefficients on the effect of heterogeneity are larger in all grades and subjects, second they are more precisely estimated.

The Effect of Heterogeneity at Different Socio-economic Environments

Table 10 presents 2SLS estimates for the two subjects in all grades of a model that allows for an interaction between heterogeneity and the SES index. The equations include other determinants (religious dummy, school size, grade enrollment, hours of instruction and locality dummy variables) but their coefficients are not reported in the table. The interaction between heterogeneity and the SES index is treated as endogenous and it is instrumented using the interaction of the 1984 values of these two variables.[17]

[17] As noted above this additional regressor was instrumented as well, but we do not present the first stage regressions of these models because they do not add more information to results of Table 3.

The sign of the Heterogeneity*SES interaction term is positive in all six, but one, columns in Table 10. However, only in two cases the standard errors are small enough to make the effect not significantly different form zero at the 5 percent significance level.

It seems, therefore, that the heterogeneity negative effect is smaller in schools with low average SES, but at the mean of the data of our sample the effect is still negative. Actually the sign reverses in the range of the SES index observed in the data only in two cases, 5th grade math and 3rd grade reading. In the first cases the reversal in sign occurs for schools in the upper quintile of the SES index distribution, in the second case it involves less than ten percent of the schools, those at the bottom end of the SES distribution.

The Effect of Heterogeneity in Small and Large Localities

Since much of the administrative effort to enhance integration in primary schools was initiated in large cities, primarily Tel-Aviv and Jerusalem, we examined the sensitivity of our results to restricting the sample to schools in municipalities that have at least 10 schools. This sample includes schools from Tel-Aviv, Jerusalem, Haifa, Cholon, Beer-Sheva, Bat-Yam, and about 85 more schools from six other cities.[18]

Estimation of a model with heterogeneity, SES and their interaction, all instrumented, yields results that are qualitatively very similar to those reported for the full sample.

There are two points to note with regard to these estimates. First, the IV estimates are almost four times larger than the OLS estimates, suggesting that the bias in neglecting to account for

[18]The sample could not be limited to just the three largest cities in Israel because such a sample could not provide enough variation in the political instruments used to estimate the effect of hours of instruction.

the endogeneity of class-heterogeneity is substantial. Second, the 'metropolitan' sample yields class-heterogeneity effects that are only 20 percent larger than those obtained with the full sample.

Based on these estimates we concluded that the relationship between class-heterogeneity and scholastic achievements does not vary substantially between large localities and the overall sample. To further substantiate this conclusion we estimated our basic model with a sample that included only the smallest localities, which contribute no more than five schools to our sample. Using the 4th graders results we find that indeed, the effect of class-heterogeneity on achievement is smaller and less precisely estimated than that obtained from the 'metropolitan' sample, or, for that matter, than the effect reported earlier for the full sample. The relative imprecision of these estimates is due not only to a smaller sample, but also to the imprecise identification of the effect of weekly hours of instruction per student on achievement. We have seen earlier that not instrumenting at all for weekly hours of instruction leads to a downward bias in the effect of class-heterogeneity on achievement, and this is most likely why the 'small localities' sample yields a very small class-heterogeneity effect on achievement. Another reason for the lower estimated effect of heterogeneity in the smaller localities sample is the level of heterogeneity in its schools, which is approximately 10 percent lower than in the 'metropolitan' sample.

It is worth noting that when large municipalities are excluded, the OLS and 2SLS estimates of the effect of class-heterogeneity on schooling outcomes are not significantly different, as they are in the sample of large cities. More specifically, the OLS effects are negative, as are the 2SLS point estimates, with relatively precise point estimates. This is an important result since it should be recalled that in the large cities the level of integration tends to be the result of school zoning policy, while in the smaller municipalities the extent of integration is related primarily to the natural distribution of the student population. This stratification is a 'crude' comparison between a treatment group (the large municipalities) and a quasi-non-random control group.

7. Interpretation of the Results

A central result of this study is that the estimated effect of class-heterogeneity on schooling outcome is very sensitive to the specification of the outcome equation and the handling of endogeneity of certain explanatory variables. The identification of this effect is somewhat complex since class-heterogeneity is directly affected by the policy decisions of local school authorities, and is also used by the Ministry of Education in conjunction with other school characteristics to determine resource allocation. Since these two mechanisms involve two different levels of decision making, which are often not synchronized, it is probably common to have schools which are targeted by the local authority to become more integrative and at the same time are not targeted by the central or local school authority to receive more hours of instruction and larger class-size, so as to increase the weekly hours of instruction per student. This is ironic since one of the most important arguments raised in favor of educational integration is a more equal distribution of educational inputs, based on the perception that low achieving schools are also 'poor' schools in terms of educational resources. It is therefore claimed that the integration of weak students in established schools will increase the equality of input distribution and provide a more efficient learning environment for weak students. The evidence presented above suggest that the more integrated schools do get a higher allocation of hours of instruction, but classes in these schools are not any smaller than in other schools.

Our results suggest that an integration policy that leads to higher school-heterogeneity will reduce the efficiency of the learning environment. In theory the learning externalities that result from exposing weak students to strong students may be large enough to offset the efficiency losses. However, all the studies in Israel that concentrated on measuring the externality effect found a positive but very weak effect, with some studies showing an effect not significantly different from zero. However, as already noted earlier, none of these studies (except for the few experimental ones) treated the measures of integration as policy variables that needed to be properly identified in estimation.

In this study we are unable to identify the peer effect although we can estimate its upper bound. The coefficient of the SES index is this upper bound as seen from equation (2). This coefficient is always negative (recall that higher SES index indicates lower socioeconomic status), indicating that the upper bound for the peer effect is positive. To compare this effect to that of class-heterogeneity we resort to input/output elasticities. The negative class-heterogeneity elasticity is always higher than the positive upper bound of the peer effect. Using for example the 4th grade math results of Table 6, column 4, yields at the mean of the sample a class-heterogeneity elasticity of -.110 [=(-37.6*.202)/68.8] while the upper bound peer elasticity of the peer effect is .078. Using the 5th grade math results of Table 7, column 5, yields elasticities of -.146 and .097. Clearly the net effect of school integration on math test scores is negative. The results for reading scores suggest the efficiency effect is offset (4th grade) or more than offset (5th grade) by the upper bound of the externality effect. The 3rd grade elasticities, based on the estimates of Table 5, indicate that the negative efficiency effect is always greater in absolute value than the upper bound to the positive externalities effect. However, the 3rd grade elasticities are generally lower than those of the 4th and 5th grades. For example, the 3rd grade reading efficiency and externality elasticities are -.064 and .026, respectively.

How large is the effect of class-heterogeneity on academic achievement? This issue can be analyzed by comparing the effect of class-heterogeneity to that of weekly hours of instruction per student. We evaluate the relative effects based on a comparison of the input-output elasticities which are based on the estimates in columns (4) and (10) of Tables 6-7 and evaluated at the mean of the 4th and 5th graders samples. The 4th grade class-heterogeneity elasticities are -.110 for math and -.097 for reading. The 5th graders elasticities are very similar, being -.051 and -.146, respectively. The weekly hours of instruction (per student) elasticity is higher: for 4th graders it is .112 for math and .142 for reading. For 5th graders it is .159 and .228, respectively. The class size elasticity is -.067 and -.099 for 4th graders for reading and math, respectively. For 5th graders it is -.159 and -.123, respectively.

First we should note that these elasticities are not trivial. They indicate that a 10 percent reduction in class heterogeneity leads to an increase of about 1-1.5 percent in math and reading test scores. In order to achieve the same gain in test scores via a change in school inputs, hours of instruction per student will have to increase by about 6 percent, or by a 15 percent reduction in class size.

An alternative assessment of the magnitude of these effects can be based on the notion of "effect size" discussed in the literature on the effects of school inputs on cognitive achievement. The "effect size" is a summary statistic defined as the ratio of the test score change resulting from a given change in an input to the standard deviation of the empirical score distribution.[19] We estimate the effect size of lowering class-heterogeneity by 20. It is interesting to compare the effect size of class-heterogeneity with the effect size of the inputs computed also for a 20 percent decline in class size and in hours of instruction per student. Using first the IV estimates from columns 2 and 6 in Table 6, we obtain an effect size for class heterogeneity that ranges from $.055\sigma$ (reading, 5th grade) to $.224\sigma$ (math, 5th grade). The respective class size effect size range from $.068\sigma$ (math, 4th grade) to $.201\sigma$ (math, 5th grade). Thus the highest effect size estimate for heterogeneity and class size changes are about equal. Repeating the same computations for heterogeneity and hours per student, using the estimates from columns 4 and 8 of Table 6, suggest larger effect size estimates for hours per student than for class heterogeneity: the class heterogeneity effect size ranges from $.113\sigma$ (reading, 5th grade) to $.234\sigma$ (math, 5th grade), the respective hours per student effect size ranges from $.175\sigma$ (math, 4th grade) to $.320\sigma$ (math, 5th grade). The ratio of the effect size of class heterogeneity relative to the effect size of hours per student ranges from a third to two thirds. However, as noted already by Angrist and Lavy (1996), even apparently small effect sizes can translate into large

[19] Finn and Achilles (1990) discuss two variations on this measure, one using the standard deviation of test scores among students and one using the standard deviation of class means. Since the overall variance is naturally larger than the between class variance, measures based on the first standard deviation are always smaller than measures based on the second. We are limited to reporting results based on the second measure because we do not have the micro score data for 4th and 5th graders. Note, however, that since class-heterogeneity is a class-level intervention, it seems reasonable to measure impact relative to the distribution of average class outcomes. Moreover, in a world of constant additive treatment effects with homoscedastic residual variation, changes in class-heterogeneity have no effect on the within-class distribution of scores.

movements through the score distribution. For example, the gap between the lower and upper quartiles of the 5th grade reading scores distribution is just 10 points and the gap between the median and the third quartiles is only 4.4 points.

8. Conclusions

The effect of school integration on academic achievement is explained via two main effects. The first is the 'peer' effect, namely externalities that are induced by the composition of teaching and learning environment. The second is the efficiency effect which reflects the reduced ability of the teacher to teach and of the pupil to learn in a heterogeneous environment. While most previous studies in Israel concentrated on measuring the externality effect, our study focused on the efficiency effect and provided an upper bound estimate to the peer effect . We presented a variety of OLS and IV estimates of the effect of class-heterogeneity on the achievement of Israeli elementary school children. The raw large negative correlation between achievement and class-heterogeneity is clearly a result of the association between more heterogeneous classes and lower SES among pupils. The effect of class heterogeneity is also very sensitive the control for school size and educational resources such as hours of instruction and class size.

Even after controlling for all these important variables the estimated effect of class-heterogeneity is significantly biased downward since the extent of school integration is a policy variable that is not independent of unmeasured determinants of cognitive achievement. IV estimates using the lagged values of class-heterogeneity and class SES consistently show a negative association between class-heterogeneity and student achievement. The estimated efficiency effect (of class-heterogeneity) strongly dominates the estimated peer effect (upper bound). These effects are apparent for both the math and reading scores of 3rd, 4th and 5th graders. The effects are very different from those reported in earlier observational studies of school integration but consistent with experimental studies of school integration in Israel.

The magnitudes of the effects are significant relative to the math and reading comprehension score distributions. Several Israeli local authorities recently began implementing a new policy that will allow parents more freedom in choosing an elementary school for their child instead of the current strictly enforced school zone rules. Our results suggest that the increased school homogeneity that will necessarily follow from this liberalization in the education system will have a positive effect in terms of schooling efficiency and the implied gains in scholastic achievement. This gain could lead to impressive improvements in Israeli test scores although this must be weighed against the setback in the efforts to achieve social integration through the education system. However, it should be noted that the negative effect of class heterogeneity can be offset by a relatively modest increase in hours of instruction or by a reduction in class size.

Data Appendix

The data used in this study came from a number of Ministry of Education files containing administrative records and test results. Since these files are not maintained for research purposes, we had to link information from a variety of sources, and to impute information that was incomplete in one source with information from other sources.

1991 data (4th and 5th graders)

A computerized data file received from the Central Bureau of Statistics (1991) survey of schools included 2073 5th grade classes in 1,027 Jewish public (secular and religious) schools in 2,073 (non-special education).[20] These data contained information on weekly hours of instruction per student. Data on average test scores were provided in two forms: Ministry of Education programmers provided us with one file containing information on average test scores and numbers of test takers for 1,733 of the classes (about 85 percent) and we were also provided with a file that contained average test scores and numbers of test-takers for each grade in each school for 1,978 of the classes. Among the 296 classes missing class-level average scores, school-level averages were available for all but 5. Since there was never more than one class missing a class-level score, and we knew the number of test-takers in each school and in each class with non-missing scores, we were able to impute the missing class-level average for all but the 5 classes missing both class-level and school-level averages. Finally, the school size, SES index and its standard deviation were added to the linked and imputed class/school data set from a separate Ministry of Education file on schools. The 1990 SES index and its standard were available for every school in the database, but their 1984 values were missing for few schools, mainly but not only new schools.

The construction of the 4th graders data set was similar to that of the fifth graders. A computerized file from the Central Bureau of Statistics (1991) survey of schools included 2,106 4th grade (non-special education) classes in 1,039 Jewish public schools. We were provided with

[20]The published Central Bureau of Statistics (1991, p. 67) report indicates that there were 1,081 Jewish public elementary schools in 1990/91, although not all of these had regular (non-special education) classes and not all had enrollment in all grades.

class-level average scores for 1,769 of the 2,059 classes and school-level averages for 2,025 of the 2,059 classes. Among the 290 classes missing class-level average scores, school-level averages were available for all but 4. Since there was never more than one class missing a class-level score, and we knew the number of test-takers in each school and in each class with non-missing scores, we were able to impute the missing class-level average for all but 4 of the classes missing both class-level and school-level averages. The SES index, its standard deviation and school size were then added as with the 5th graders.

We checked the imputation of class-level averages from school averages by comparing the school and class averages in schools with one class and by comparing the imputed and non-imputed data. School and class-level averages matched almost perfectly in schools with one class. We were unable to detect any systematic differences between schools that were missing some class-level data and schools that were not. The empirical findings were not sensitive to the exclusion of the imputed class-level averages.

1992 data (3rd graders)

Construction of the 3rd graders data set differed from the construction of the 4th and 5th graders data sets since we were provided with micro data on the test scores of 3rd grade pupils. As with the 4th and 5th graders, we began with the Central Bureau of Statistics (1993) survey of schools. This includes 1,042 Jewish public schools with 3rd grade pupils, in 2,193 (non-special education) classes.[21] Data on weekly hours of instruction per student, provided by the

[21]The published Central Bureau of Statistic (1993, p. 125) report indicates that there were a total of 1,080 Jewish public elementary schools in 1991/2.

Ministry of Education, contained records with information on hours of instruction for 2,162 of these classes.

We used micro data on the test scores of 3rd graders to compute average math and reading scores for each class. Score data were available for 2,144 of the 2,162 classes with class size information obtained from the CBS survey of schools. Finally, we added information on the SES index, its standard deviation and school size from a Ministry of Education file containing information on schools. There was no information on the SES index for 34 of the 2,144 classes with data on size and test scores, resulting in a 3rd grade base line sample of 2,111. This is probably because new schools will not have had an SES index assigned at the time data in our school-level file was entered into the record-keeping system. The sample with a full data set, including the class-heterogeneity measure, included 963 classes.

References

Algrabi, M., "Indices to Characterize a School's Social Composition and a Method for Allocating Additional Funds Between Schools," (Hebrew), *Megamot* 21 (1975), 219-227.

Alperovich, G., "The Economics of Choice in the allocation of Intergovernmental Grants to Local Authorities," *Public Choice* 44 (1984), 285-286.

Angrist, J.D., "Grouped-Data Estimation and Testing in Simple Labor Supply Models," *Journal of Econometrics* 47 (1991), 243-266.

Angrist, J.D., and Guido Imbens, "Average Causal Response in Models with Variable Treatment Intensity," *Journal of the American Statistical Association* 90 (1995), June.

Angrist, J.D., and Victor Lavy, "Using Maimonides' Rule to Estimate the Effect of Class Size on Children's Academic Achievement," Draft, Hebrew University of Jerusalem, Department of Economics, June 1996.

Alexander, K.L., and M., Cook, "Curricula an Course Work: A surprise Ending to a Familiar Story," *American Sociological Review* 47 (1982): 626-640.

Arnot, R. and J. Rowse, "Peer Group Effects and Educational Attainment," *Journal of Public Economics* 32 (1987), 287-305.

Arzi, Y., and Y. Amir, "Intellectual and Academic Achievements and Adjustment of Underprivileged Children in Homogeneous and Heterogenous Classrooms," *Child Development* 348 (1977), 726-729.

Barr, R., R, Dreeben, and N., Wiratchai, How School Work, Chicago: University of Chicago Press

Ben-Ari, R., S, Sharan and A., Amir, "Ethnic Mixing in Education," In Amir, Sharan and Ben-Ari (Eds.) Integration in Education, 1985:15-31, Tel-Aviv, Am Oved.

Behrman, Jere, and Nancy Birdsall, "The Quality of Schooling Alone is Misleading," *American Economic Review*, 73 (1993), 5, 928-946.

Besley, Timothy, and Anne Case, "Unnatural Experiments? Estimating The Incidence of Endogenous Policies," NBER Working Paper No. 4956 (1994).

Bradley, L.A., an G.W., Bradley, "The Academic Achievement of Black Students in Desegregated Schools: A Critical Review," *Review of Educational research*, 47 (1977), 399-499.

Betts, Julian R., "Does School Quality Matter? Evidence from the National Longitudinal Survey of Youth, " *Review of Economics and Statistics* (forthcoming).

Card, David, and Alan Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100 (1992a), 1-40.

Card, David, and Alan Krueger, "School Quality and Black-White Relative Earnings: A Direct Assessment," *Quarterly Journal of Economics* 107 (February 1992).

Card, David, and Alan krueger, "Labor Market Effects of School Quality: Theory and Evidence," NBER Working Paper No. 5450, February 1996.

Central Bureau of Statistics, *Educational Institutions: Kindergartens, Schools, Post-Secondary and Higher Education 1990/91*, Series of Education and Culture Statistics, No. 198, reprinted from the *Supplement to the Monthly Bulletin of Statistics*, No. 9, 1991.

Central Bureau of Statistics, *Educational Institutions: Kindergartens, Schools, Post-Secondary and Higher Education 1991/92*, in the *Supplement to the Monthly Bulletin of Statistics*, No. 11, November 1993 (Hebrew only).

Central Bureau of Statistics, *Educational Institutions: Kindergartens, Schools, Post-Secondary and Higher Education 1992/93*, Series of Education and Culture Statistics, No. 219, reprinted from the *Supplement to the Monthly Bulletin of Statistics*, No. 6, 1994.

Chamberlain, G., "Panel Data." In *Handbook of Econometrics*, Editors: Z. Griliches and M.D. Intriligator. Amsterdam: North Holland.

Chen, M., and D. Kfir (1981). "How Did Middle Schools Led to Student's Progress," *Studies in Education*, 32, pp. 59-82 (Hebrew).

Chen, M., A., Lewy, an Ch., Adler, Educational Process and Outcome: Evaluation of the Junior-High School Reform. Jerusalem: Ministry of Education (1978), (Hebrew).

Chen, M., and D. Kfir (1981). "How Did Middle Schools Led to Student's Progress," Studies in

Coleman, J.S., et al, *Equality of Educational Opportunity*, Washington, DC: US GPO, 1966. Resh, Classroom Composition and Pupil Achievement: A study of the Effect of Ability-based Classes, (1986) New York: Gordon & Breach.

Dar, Y., and N. Resh, Homogeneity and Heterogeneity in Education: Interaction between Personal Resources and the Learning Environment in the Effect of Scholastic Achievement," (1981), The Hebrew University of jerusalem.

Dar, Y., and N. Resh, Educational Integration and Learning Achievement: A Summary and Evaluation of The Research in Israel," Megamot 31 (1988): 180-207, (Hebrew).

Farley, and W., Allen, The Color Line an the Quality of Life in America. New York: Russell Sage foundations (1987).

Gerard, H.B., and N., Miller, School Desegregation. (1975) New York: Plenum.

Good, T., and S., Brophy, Looking in Classrooms. New York: Harper and Row, 1987.

Card, David, and Alan Krueger, "Does School Quality Matter? Returns to Education and the Characteristics of Public Schools in the United States," *Journal of Political Economy* 100 (1992a, 1-40.

Hanushek, Eric, "The Economics of Schooling: Production and Efficiency in Public Schools," *Journal of Economic Literature* 24 (September 1986), 1141-1177.

Hanushek, Eric, "School Resources and Student Performance," in *Does Money Matter? The Effect of School Resources on Student Achievement and Adult Success*, edited by Gary Burtless, Brookings Institution: Washington, DC, 1996.

Holtz-Eakin, D., W., Newey and A. S., Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica* 56 (1988): 1371-1395.

Hedges, L.V., R.D., Laine, and R. Greenwald, "Does Money Matter? A Meta-analysis of Studies of Analysis of Studies of the effects of Differential School Inputs on Student Outcomes," *Educational Researcher* 23 (1994), 5-14.

Katz. Yaacov and A. Ben-Yochanan (1988). "Academic Achievement as a Function of Integration in an Israeli Elementary School." *British Educational research Journal*, vol. 14, No 3, pp: 313-320.

Klinov, R, "The Response of the School System to the 1980-86 Budget Cuts Schools," in Public Services Under Stress: A Canadian-Israeli Policy Review (1993), 206-227.

Klein, W., and Y. Eshel, Integrating Jerusalem Schools. (1980) New York: Academic Press.

Lavy, Victor, "Endogenous School Resources and Cognitive Achievement in Primary Schools in Israel," Hebrew University, Falk Institute Discussion Paper No. 95.03, 1995.

Lavy, Victor, "Endogenous School Resources, Local Government Political Affiliation, and Scholastic Achievement in Israel," Hebrew University, Draft, May 1996.

Lee, V.E. and A.S. Bryk, "Curriculum Tracking as Mediating the Social Distribution of High Schools," *Sociology of Education,* 61 (1988),78-94.

Lewy, A., "Class composition, the Gap Between Ethnic Groups and Progress in Achievements," (1977) *Megamot* 23:88-97 (Hebrew).

Lewy, A., and M. Chen, "Closing or Widening the Achievement Gap," (1974) Research Report No. 6. Tel-Aviv University: School of Education.

Link, C., and J. Mulligan, "Classmates' Effect on Black Student Achievements in Public Schools Classrooms," *Economics of Education Review* 10 (1991), 297-310.

Loeb, Susanna, and John Bound, "The Effect of Measured School Inputs on Academic Achievement: Evidence From the 1920s, 1930s, and 1940s Birth Cohorts," NBER Working Paper No. 5331, November 1995.

Minkevich, A., A. Davis, and D. Bashi (1980). Educational Achievements in Primary Schools in Israel, Jerusalem, Magnum Press (Hebrew).

Moulton, B.R., "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics* 32 (1986), 385-397.

Mueller, D., Chase, C.I., and J.D. Walden, "Effects of Reduced Class Size in Primary Classes," *Educational Leadership* 45 (1988), 48-50.

National Center for Education Feedback, "*The Results of Reading and Mathematics Achievement Tests Given to Fourth and Fifth Graders in June 1991*, (Hebrew pamphlet), Jerusalem: Israel Department of Education, October 1991.

National Center for Education Feedback, *Selected Analyses of the Results of the June 1992 Third Grade Achievement Tests*, (Hebrew Pamphlet), Jerusalem: Israel Department of Education, March 1993. Rosenzweig, M., and K. Wolpin, "Evaluating the Effects of Optimally Distributed Public Programs: Child Health and Family Planning Interventions." *American Economic Review*, 76 (1986):470-82.

Rosevitch, S. and A. Weiss, "Beneficiaries From Federal Transfers to Municipalities: The Case of Israel." *Public Choice* 76 (1993): 335-346.

Slavin, R., Ability Grouping an Student Achievement in Elementary Schools: A Best Evidence analysis," (1986) unpublished manuscript, Baltimore: Center for Effective elementary and Middle Schools. The Johns Hopkins University.

Spady, W.G. "The Impact of School Resources on Students," (1973) in Kerlinger, F.N. (ed.), Review of Research in Education, Itacha, IL: Peacock.

Steinberg Bernard. (1990). "education and Integration in israel: The First Twenty Years.

Summers, A. and B. Wolfe, "Do Schools Make a Difference?" *American Economic Review* 67, (No. 4, September 1977): 639-652.

Svirski, S. (1990). Education in Israel, Tel-Aviv, Brerot (Hebrew).

________, "Separating and Mixing Students for Learning: Concepts and Research." *Pedagogish Tijdchrift* 19. No. 2, (1994):109-126.

Smilanski, S. and L. Shfatia, "Socio-cultural Integration and Other Characteristics if Human Environment in the Classroom as Related to Achievement in Grades One and Two," (1977) Jerusalem: Szold Institute.

Smilanski, S. and L. Shfatia, "The Relationship Between Integration, Class Level Variables and Achievements in First and Second Grade," *Megamot*, 23 (4): 79-87 (Hebrew).

Speakman, R. and F. Welch (1995). "Does School Quality Matter? -- A Reassessment." Texas A&M University (mimeo.).

Tel-Aviv Municipality. (1976). "Achievement, Student's Background, and Teachers's Characteristics in Primary Schools," Special Publication, The Department for Research and Statistics (Hebrew).

Yogev, A., and C. Ayalon (1985). The Law of Free High School and Equality of Opportunities in High School Education," Tel-Aviv University, The Sapir Institute (Hebrew).

Yogev, A., and D. Kfir. (1981). "Factors Affecting High School Schooling Choice," Megamot, 27: 139-153 (Hebrew).

Table 1: Descriptive Statistics

Variable	Mean	S.D.	Quantiles 0.10	0.25	0.50	0.75	0.90
$Heterogeneity_{90}$	.23	.09	.03	.16	.21	.26	.31
$SES\text{-}Index_{90}$	.17	.17	.032	.05	.11	.22	.41
$Heterogeneity_{85}$	.45	.12	.32	.40	.46	.51	.57
$SES\text{-}Index_{85}$	.23	.22	.03	.07	.15	.33	.56
Hours of instruction	1.46	.41	1.07	1.17	1.36	1.64	2.00
class size, 4th grade	30.32	6.39	22	26	31	35	38
class size, 5th grade	29.95	6.60	21	26	31	35	38
class size, 3rd grade	30.48	6.23	22	26	31	35	38
cohort size, 4th grade	75.41	36.71	29	49	70	98	121
cohort size, 5th grade	74.193	37.23	28	48	69	96	123
cohort size, 3rd grade	76.81	36.15	32	50	72	100	125
Religious	.76	.43					
School size	503.50	232.94	207	333	491	648	800
Reading, 4th grade	72.47	8.012	62.40	67.67	73.32	78.2	82.00
Reading, 5th grade	74.38	7.68	64.15	69.83	78.39	79.84	83.32
Reading, 3rd grade	86.13	6.12	78.40	82.99	87.24	90.74	93.10
Math, 4th grade	68.841	8.79	57.50	63.57	69.32	74.96	79.38
Math, 5th grade	67.30	9.59	54.87	61.23	67.78	74.09	79.40
Math, 3rd grade	84.10	6.82	75.00	80.24	84.76	89.01	91.92

Notes:
sample sizes:
4th grade: 2018 classes, 1030 schools
5th grade: 1989 classes, 1018 schools
3rd grade: 2068 classes, 1040 schools

Table 2: OLS estimates

	5th Grade								4th Grade							
	Reading comprehension				Math				Reading comprehension				Math			
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16
Mean Score (S.D.)		74.3 (7.4)				67.3 (9.3)				72.4 (7.6)				68.8 (8.6)		
Heterogeneity	-31.205 (1.829)	- 6.711 (1.982)	-7.269 (1.957)	- 4.737 (2.227)	-28.877 (2.322)	-1.780 (2.589)	-.2366 (2.587)	-4.289 (2.845)	-27.798 (1.838)	-1.782 (1.981)	-2.300 (1.958)	.532 (2.150)	-22.476 (2.015)	2.904 (2.310)	2.349 (2.311)	.08 (2.4
SES-Index		-24.101 (1.094)	-28.312 (1.225)	-24.986 (1.656)		-.26663 (1.429)	-27.038 (1.620)	-26.163 (2.118)		-25.334 (1.091)	-29.552 (1.211)	-25.652 (1.587)		-24.715 (1.273)	-25.474 (1.429)	-24.4 (1.8
Religious			3.101 (.409)	3.748 (.452)			1.390 (.541)	2.251 (.578)			2.909 (.406)	3.435 (.437)			1.377 (.479)	2.0 (.50
School size			.001 (.001)	.005 (.001)			.003 (.001)	.005 (.001)			.001 (.0007)	.005 (.0009)			.002 (.0008)	0.0 (.0
Town effects				√				√				√				√
N		1996				1996				2028				2028		

Notes: Unit of observation is the average score in the class, standard errors in parenthesis.
Heterogeneity and SES-Index co-efficients are divided by 100.

Table 3: Reduced-form Estimates

	4th Graders				5th Graders			
	Reading	Math	Hetero-geneity	SES-Index	Reading	Math	Hetero-geneity	SES-Index
Regressor	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Heterogeneity_{84}$	-3.365 (1.591)	-3.872 (1.820)	.117 (.018)	.006 (.022)	-.794 (1.729)	-3.716 (2.187)	.109 (.011)	.004 (.024)
$SES\text{-}Index_{84}$	-12.210 (.961)	-9.954 (1.099)	.095 (.0111)	.289 (.014)	13.095 (1.000)	-11.200 (1.265)	.097 (.011)	.287 (.014)
Religious	1.152 (.397)	.176 (.454)	.036 (.004)	.112 (.006)	1.396 (.413)	-.240 (.522)	.033 (.005)	.108 (.006)
School Size	.006 (.0064)	.007 (.001)	-.00005 (.00001)	-.0001 (.00001)	.007 (.001)	.008 (.001)	-.00005 (.00001)	-.00012 (.00001)
Town Effects	√	√	√	√	√	√	√	√
N	1933		1933	1933	1910		1910	1910

Notes: Heterogeneity = The standard deviation of the 1990 SES-Index distribution
SES-index = The 1990 SES-Index

Table 4: 2SLS Estimates, 4th and 5th grade

	Reading comprehension				Math			
	4th grade		5th grade		4th grade		5th grade	
Regressor	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Heterogeneity	.739	-26.982	-5.221	-8.661	.319	-31.852	-4.740	-34.971
	(2.029)	(14.366)	(2.092)	(15.711)	(2.326)	(16.180)	(2.659)	(19.925)
SES-Index	-26.473	-33.377	-25.179	-42.657	-25.144	23.957	-26.294	-27.196
	(1.505)	(6.557)	(1.564?)	(7.008)	(1.726)	(7.389)	(1.984)	(8.888)
Religious dummy	3.481	5.845	3.734	6.273	2.143	3.637	2.294	3.837
	(.417)	(.616)	(.431)	(.618)	(.478)	(.694)	(.547)	(.784)
School size	.004	.001	.105	.0004	.005	.002	.005	.003
	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)	(.001)
Town effects	√	√	√	√	√	√	√	√
N	1983		1910		1933		1910	

Notes: Unit of observation is the average score in the class, standard errors in parenthesis.
All regressions include town fixed effects.

Table 5: 2SLS Estimates, 3rd grade

Regressor	Reading comprehension				Math			
	(1) OLS	(2) OLS	(3) OLS	(4) IV	(5) OLS	(6) OLS	(7) OLS	(8) IV
Heterogeneity	-12.938 (1.528)	-5.413 (1.815)	-1.884 (1.728)	-27.334 (14.940)	-15.170 (1.362)	-2.996 (1.567)	-4.999 (1.959)	-24.978 (16.263)
SES-Index	-	-7.723 (1.036)	-11.065 (1.297)	13.734 (6.517)	-	-12.497 (.895)	-5.666 (1.471)	-4.578 (7.085)
Religious	-	-	1.172 (.353)	2.738 (.537)	-	-	.045 (.400)	-.768 (.585)
School size	-	-	.002 (.0007)	-.0005 (.001)	-	-	.003 (.001)	.001 (.001)
Heter*SES index	-	-	-	-	-	-	-	-
Town effects			√	√			√	√
N		2100		2000		2100		2000

Notes: Unit of observation is the average score in the class, standard errors in parenthesis.

Table 6: Heterogeneity and School Resources: 4th grade

Regressor	Reading comprehension				Math			
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Heterogeneity	.778	-25.370	1.211	-34.619	.387	-30.655	.894	-37.584
	(2.031)	(14.254)	(2.032)	(16.742)	(2.329)	(16.009)	(2.329)	(18.227)
SES-Index	-26.307	-35.288	-23.992	-44.577	-24.952	-25.354	-22.251	-32.347
	(1.511)	(6.510)	(1.681)	(8.5551)	(1.733)	(7.356)	(1.927)	(9.309)
Class size	.036	-.131	-	-	.040	-.098	-	-
	(.028)	(.061)			(.032)	(.069)		
Hours per student	-	-	-2.042	7.030	-	-	-2.375	5.267
			(.620)	(3.530)			(.711)	(3.843)
N	1933				1933			

Notes: Unit of observation is the average score in the class, standard errors in parenthesis.
The following variables are also included as regressors in all the regressions reported above: town dummies, school size, cohort size, and a religious dummy.

Table 7: Heterogeneity and School Resources: 5th grade

Regressor	Reading comprehension				Math			
	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV	(7) OLS	(8) IV
Heterogeneity	-5.138 (2.086)	-10.458 (16.526)	-4.489 (2.076)	-18.975 (19.241)	-4.675 (2.657)	-37.500 (20.918)	-3.985 (2.648)	-48.681 (24.649)
SES-Index	-24.314 (1.574)	45.787 (7.296)	-21.911 (1.637)	52.139 (8.529)	-25.489 (2.005)	-31.119 (9.236)	-22.820 (2.088)	-39.465 (10.926)
Class size	.115 (.030)	-.245 (.088)	-	-	.106 (.038)	-.320 (.111)	-	-
Hours per student	-	-	-2.696 (.438)	8.026 (3.164)	-	-	-2.858 (.559)	10.512 (4.053)
N	1910				1910			

Notes: Unit of observation is the average score in the class, standard errors in parenthesis.

71

72

Table 8: Summary of the effect of class heterogeneity

		IV no inputs	IV with class size	IV with hours per student
		(1)	(2)	(3)
4th grade				
	Math	-31.852 (16.80)	-30.655 (16.09)	-37.584 (11.227)
	Reading comprehension	-26.982 (14.366)	-25.370 (14.254)	-34.619 (16.742)
5th grade				
	Math	-34.971 (19.925)	-37.500 (20.918)	-48.681 (24.649)
	Reading Comprehension	-8.661 (15.771	-10.458 (16.526)	-18.975 (19.241)

Note: standard errors in parenthesis.

Table 9: Summary of the effect of class heterogeneity: restricted sample

		IV no inputs	IV with class size	IV with hours per student
		(1)	(2)	(3)
4th grade				
	Math	-45.227 (12.722)	-46.625 (12.790)	-52.748 (14.174)
	Reading comprehension	-49.054 (11.330)	-50.432 (11.461)	-58.486 (13.216)
5th grade				
	Math	-33.547 (14.83)	-40.058 (15.662)	-44.848 (18.361)
	Reading Comprehension	-25.279 (12.131)	-29.357 (12.702)	-32.359 (14.427)

Note: standard errors in parenthesis.

76

Table 10: Additional IV Results, Heterogeneity and SES Interacted

	Reading Comprehension			Math		
	4th Grade	5th Grade	3rd Grade	4th Grade	5th Grade	3rd Grade
	(1)	(2)	(3)	(4)	(5)	(6)
Heterogeneity	-34.846 (16.737)	-18.315 (17.820)	-23.579 (12.886)	-36.917 (17.159)	-46.943 (22.075)	-24.4141 (14.535)
SES-Index	-42.647 (10.357)	-66.869 (10.468)	-25.384 (8.989)	-38.037 (10.618)	-18.281 (12.968)	-9.924 (10.139)
Heter*SES	-10.811 (50.650)	64.060 (45.873)	64.711 (32.285)	31.868 (51.926)	168.820 (56.828)	25.599 (36.416)

Note: standard errors in parenthesis.

מכון מילקן לחקר מערכות חינוך
ע מ ו ת ה

השפעת ההטרוגניות של אוכלוסיית התלמידים בבית הספר על הישגים לימודיים בחינוך היסודי בישראל

פרופ' ויקטור לביא

ירושלים, סיון תשנ"ז, יוני 1997

פרופ' דניאל י. אלעזר
מנהל אקדמי

צבי ר. מרום
מנהל כללי

פרוייקט משותף של המרכז הירושלמי לענייני ציבור ומדינה וקרן משפחת מילקן-לשעבר אייסס

בית מילקן, רח' תל-חי 13, ירושלים 92107 טל. 02-5619281 פקס. 02-5619112
דואר אלקטרוני: E-Mail: elazar@vms.huji.ac.il

מכון מילקן לחקר מערכות חינוך (לשעבר אייסס) - הוקם במרכז הירושלמי לענייני ציבור ומדינה בשנת 1990 בשיתוף עם קרן משפחת מילקן מארה״ב.

מטרות המכון: ליזום, לתכנן ונהל תכניות מחקר יישומיות בתחום מערכת החינוך בישראל. המכון שואף להגיש שירותי יעוץ למקבלי ההחלטות בישראל בתחום החינוך ולסייע בעיצוב מגמותיה של מדיניות החינוך העכשווית ברוח צרכי המאה ה-21 .

חבר העמיתים

- פרופ׳ דניאל י. אלעזר, נשיא המרכז הירושלמי לענייני ציבור ומדינה
  גב׳ מתוקה בנג׳מין, מנהלת בתי הספר על יד בית הכנסת ע״ש סטפן וייז
  דר׳ אריה הכט, יועץ לנושאי שלטון מקומי ומינהל ציבורי
  דר׳ ג׳וליוס לסנר, מנכ״ל קרן משפחת מילקן
- מר לואל מילקן, נשיא קרן משפחת מילקן
- מר צבי ר. מרום, מנכ״ל המרכז הירושלמי
- פרופ׳ לואיס סולמון, דיקן לשעבר של בית הספר לחינוך באוניברסיטת UCLA ;
  נשיא מכון מילקן לתעסוקה והשקעות הון
  פרופ׳ אליעזר שבייד, האוניברסיטה העברית
- חברי הנהלה

ועדה מייעצת

ישראל פרופ׳ חיים אדלר, האוניברסיטה העברית
פרופ׳ ולטר אקרמן, אוניברסיטת בן-גוריון
פרופ׳ דן ענבר, האוניברסיטה העברית

ארה״ב פרופ׳ ג׳ימס ס. קולמן, אוניברסיטת שיקגו**
פרופ׳ ל. שולמן, אוניברסיטת סטנפורד ; נשיא לשעבר של האקדמיה הלאומית לחינוך

** נפטר

כתובת : רח׳ תל-חי 13, ירושלים 92107
טלפון : 02-5619281 פקס : 02-5619112

מחקרים ופרסומים

1. פרופ׳ דניאל י. אלעזר, **Israel's Educational System: An Introduction to a Study Program**, 1990.

2. ד״ר אריק כהן, **Education and Socio-Economic Achievements**, 1991.

3. ד״ר שמעון רוזביץ, **מגמות בתקציב החינוך בשנות השמונים**, 1993 (עברית ואנגלית).

4. פרופ׳ חיים גזיאל, **מדיניות החינוך על פרשת דרכים- בין שינוי להמשכיות: החינוך בישראל בעשור האחרון**, 1993 (עברית ואנגלית).

5. ד״ר אריק כהן, **מעורבות גורמים חיצוניים בבתי ספר**, 1993.

6. פרופ׳ דניאל י. אלעזר, **Education in a Society at a Crossroads: Implications for the Study of Israel's Education System**, 1994.

7. Dr. Jo-Ann Harrison, **Unity and Diversity of Culture and Curriculum in the Israeli Education System**, 1994.

8. ד״ר ג׳ו-אן הריסון, **אחידות ופלורליזם של תוכניות לימודים ותרבויות בית-ספריות במערכת החינוך בישראל**, 1995.

9. שושנה לנגרמן, **Survey of Israeli Research in Education 1975-1995**, 1996.

10. ד״ר אריק ועינת כהן, **החינוך האפור בישראל: התל״ן בשנות ה-90**, 1996.

11. פרופ׳ דן ענבר וד״ר מאיה חושן, **ביזור סמכויות בתחום החינוך לרשויות המקומיות** (עומד להתפרסם, 1997)

השפעת ההטרוגניות של אוכלוסיית התלמידים בבית הספר על הישגים לימודיים בחינוך היסודי בישראל

תקציר

ויקטור לביא

עבודה זו בודקת את הקשר בין מידת ההטרוגניות של התלמידים בכיתה לבין הישגים לימודיים של תלמידים בבתי הספר היסודיים בישראל. מאז פרסומו של דו"ח קולמן בארה"ב (1966) אשר הצביע על כך שהגורם השני במעלה המשפיע על הישגי התלמיד, לאחר יכולתו האישית, הוא יכולתם של בני כיתתו, הפכו ההטרוגניות הבית -ספרית והאינטגרציה הלימודית למוקד תשומת ליבם של הורים, אנשי חינוך וחוקרים במדעי החברה. הדיון בישראל בנושא זה גולש מעבר לקשר בין הטרוגניות בית- ספרית לבין הישגים לימודיים, והאינטגרציה החינוכית נתפסה גם ככלי להעמקת השילוב החברתי-תרבותי בין הקבוצות האתניות השונות בישראל. מדיניות האינטגרציה הלימודית נועדה להתמודד עם פערים בהישגים לימודיים וחברתיים, הקיימים בעיקר על פי הבחנה אתנית בין עולים מאסיה-אפריקה לבין עולים מאירופה-אמריקה.

מדיניות האינטגרציה הלימודית בישראל החלה בסוף שנות ה-60 עם ישום הרפורמה במערכת החינוך (1968), אשר במסגרתה הוקמה חטיבת הביניים. אזורי הרישום של חטיבות הביניים נועדו לשלב בבתי הספר תלמידים מרקע כלכלי-חברתי מגוון. אולם מערכת החינוך לא יזמה מהלך דומה להעמקת האינטגרציה בחינוך היסודי. בישובים רבים יזם השלטון המקומי שינויים בהגדרת אזורי הרישום כדי לשלב בבתי הספר אוכלוסיית תלמידים הטרוגנית יותר, המשקפת בעיקר גיוון חברתי-כלכלי רחב יותר. כמו כן, במספר ערים גדולות מסתמנת לאחרונה מגמה של הורים המעוניינים בהגמשת אזורי הרישום כדי לאפשר בחירה חופשית יותר של בית הספר היסודי. מימושו של מהלך כזה יוביל בהכרח לצמצום ההטרוגניות באוכלוסיית בית הספר ולהאטה באינטגרציה הלימודית-חברתית. עבודת מחקר זו מבקשת להגיש ממצאים אמפיריים עדכניים, אשר עשויים לסייע בהערכת השפעתה הצפוייה של סביבה לימודית הומוגנית יותר על ההישגים הלימודיים של תלמידים בבתי הספר היסודיים בישראל.

המחקר המדעי בישראל בנושא השפעת האינטגרציה, אשר הצטבר עד כה, אינו מעמיד לרשות מקבלי החלטות מסקנות ברורות והמלצות לקביעת מדיניות. דר ורש (1988) מביאים סיכום של 12 מחקרים

העוסקים בהשפעת האינטגרציה בישראל. החוקרים מצביעים על כך, שהתוצאות אינן חד-משמעיות ואף קיימות מסקנות מנוגדות לגבי יעילותה של מדיניות האינטגרציה. מחד, הממצאים העולים מניסויים מדעיים (ארזי ואמיר, 1976, קליין ואשל, 1980, כץ ובן יוחנן, 1988) מצביעים על השפעה חיובית של האינטגרציה, רק כאשר ההתערבות הייתה מלווה בתוספת של משאבים והתערבויות חינוכיות נוספות. מאידך, עבודות אשר אמדו את השפעתה של הטרוגניות בית ספרית 'טבעית' הניבו הערכות שנויות במחלוקת או סותרות לגבי השפעתה של מדיניות עירוב תלמידים. מרבית העבודות טענו כי השפעתו של הרכב התלמידים על השגי התלמידים הינה חיובית אך חלשה, ולעיתים אינה מובהקת; ראה לדוגמא, סמילנסקי ושפטיה, 1977, מינקביץ, דיוויס ובשי, 1980.

ניתן להצביע על שלושה מאפיינים משותפים בכל העבודות אשר בדקו את השפעת ההטרוגניות הטבעית על הישגי התלמידים. ראשית, מדדי ההטרוגניות הכיתתית שימשו כמשתנים מסבירים אקסוגניים, ולא נלקחה בחשבון האפשרות שמדיניות אינטגרציה מכוונת השפיעה על מדדים אלה. שנית, הניתוח בעבודות אלה התעלם מהשפעתם של משאבים כלכליים וחינוכיים על הישגי התלמידים. שלישית, כל העבודות התמקדו בהשפעת הרכב הכיתה על הישגי התלמידים (השפעות חיצוניות) מבלי לבדוק את השפעת הטרוגניות הכיתה על יעילות ההוראה והלמידה (השפעות יעילות).

עבודה זו נועדה לאמוד במשולב את השפעת היעילות וההשפעות החיצוניות על ההישגים הלימודיים בבתי הספר היסודיים בישראל. בחנו בקפידה את האפשרות שהיקף ההטרוגניות הבית-ספרית מושפע מהחלטותיהם של רשויות החינוך והורי התלמידים, עובדה המחייבת התייחסות להטרוגניות כאל משתנה מסביר אנדוגני. כמו כן מנטרל הניתוח את השפעתם של משאבים כלכליים וחינוכיים על הישגי התלמידים כדי לאמוד את ההשפעה "הנקייה" של הטרוגניות התלמידים על הישגיהם. מבחני המשוב בחשבון ובהבנת הנקרא משנת 1991-92 לכיתות ג-ה מהווים את בסיס הנתונים למחקר האמפירי.

ממצאי העבודה מציעים כי עליה בהטרוגניות הבית-ספרית מובילה לירידה גדולה ומובהקת בהישגי התלמידים בחשבון ובהבנת הנקרא בכיתות ג-ה בחינוך היסודי בישראל. ההשפעה השלילית הנובעת מן הפגיעה ביעילות ההוראה והלמידה אינה מפוצה על ידי ההשפעות החיצוניות החיוביות. ההשפעה השלילית הנקיה של ההטרוגניות קטנה יותר, ככל שאוכלוסיית התלמידים בבית הספר מאופיינת ברקע חברתי-כלכלי חלש יותר, אולם ההשפעה נותרת שלילית גם בבתי הספר ה"חלשים" ביותר. השימוש במדד ההטרוגניות הבית-ספרי כמשתנה אקסוגני, תוך התעלמות מהאפשרות ש"העירוב" בבית הספר הינו תוצאה של מדיניות מערכת החינוך או החלטות הורים, ומוביל לתוצאות שונות לחלוטין: להטרוגניות השפעה שלילית ברוב המקרים, אך בכל המקרים השפעה זו איננה שונה באופן מובהק

מאפס. מכאן ניתן להסיק שההתעלמות מהאנדוגניות של רמת ההטרוגניות הבית-ספרית עשויה להוביל למסקנות שגויות ביותר. כמו כן, אי נטרול השפעת המשאבים הבית-ספריים מוביל להטייה כלפי מטה (כ-20 אחוז) בהשפעת ההטרוגניות הבית-ספרית על הישגי התלמידים.

גודל ההשפעה של הטרוגניות התלמידים הינו משמעותי ביחס להתפלגות הציון הכיתתי הממוצע במדגם. לדוגמא, עליה של 10 אחוז בהטרוגניות בכיתות ד' מובילה לירידה של כ-11 אחוז בציון בחשבון ולירידה של כ-10 אחוז בציון בהבנת הנקרא. לצורך השוואה, האומד שאמדנו בעבודה זו להשפעת גודל הכיתה מציע שעליה של 10 אחוז בגודל הכיתה (כיתות ד') מובילה לירידה של 7 אחוז בציון הממוצע בהבנת הנקרא וירידה של 10 אחוז בציון בחשבון. מאידך ההשפעה של שעות הוראה שבועיות גבוהה פי שניים מההשפעה של הטרוגניות בית-ספרית. מכאן ניתן להסיק כי שינוי לא גדול יחסית בתשומות ובמשאבים (גודל כיתה, שעות הוראה שבועיות, הכשרת מורים) עשוי לפצות על ההשפעה השלילית של עליה ברמת ההטרוגניות הבית-ספרית. כלומר, מימושה של מדיניות אינטגרציה חינוכית, אשר מוביל בהכרח לעליה ברמת ההטרוגניות הבית-ספרית, אכן פוגע בהישגים לימודיים (ממוצע כיתתי), אך פגיעה זו ניתנת לנטרול על ידי תוספת משאבים צנועה יחסית.

חשיבותם של מימצאים אלה לקובעי המדיניות נובעת מן המסקנה, העולה מן המחקר, שרשויות החינוך צריכות לכלול כקריטריון להקצאת משאבים בית-ספריים לא רק את הממוצע של איכות התלמידים בבית הספר (מדד הטיפוח), כפי שנהוג עד כה, אלא גם את **השונות התוך בית-ספרית או כיתתית בהרכב התלמידים**. מדיניות הקצאת משאבים המשלבת את הממוצע והשונות של איכות התלמידים בבית הספר תהווה מענה הולם ותאפשר את מימושה של המטרה החברתית-תרבותית הגלומה במדיניות האינטגרציה מבלי לפגוע בהישגים הלימודיים. לבסוף, יש לציין שמימצאי העבודה הם חלקיים ולכן יש להעמיק את המחקר בנושא האינטגרציה והשפעותיה בחינוך היסודי והעל-יסודי, ולבסס את המדיניות החינוכית והקצאת המשאבים הלאומיים לחינוך על מידע אמין ועדכני.